The Dissertation Committee for Haibo Li Certifies that this is the approved version of the following dissertation:

Biochemical production by *Yarrowia lipolytica* from Xylose

Committee:

Hal S Alper, Supervisor

Jeffrey E Barrick

Lydia M Contreras

Andreas T Matouschek

Jennifer A Maynard

Biochemical Production by *Yarrowia lipolytica* from Xylose

by

Haibo Li

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2019

Dedication

For everyone I love

and xylose.

Acknowledgements

I would like to first thank Dr. Hal Alper for his guidance, advice and support. He interviewed me when I was applying to University of Texas at Austin, and has been a role model for me as a scientist. He tolerated my horrible writing skill and spent numerous hours correcting my papers. His knowledge and intelligence helped me went through many obstacles and he pushed me past periods of confusion and when I feel lost. He showed me what it takes to become a good scientist.

I would also want to thank my committee members. I would like to thank Dr. Lydia Contreras for her care and support. Her lab was the first lab I rotated and I have also taken a course she teaches. Both of them were very helpful for my graduate school experience and I sometime feel like she's my second advisor. I'd also like to thank Dr. Jeffery Barrick, Dr. Andreas Matouschek and Dr. Jennifer Maynard for their thoughtful questions and advices on my projects.

I want to thank all the awesome members in the Alper lab. I'd like to thank Dr. Eric Young and Dr. Sun-mi Lee for their early work on transporter engineering, and thank Dr. Leqian Liu for the early work of xylose pathway in *Yarrowia lipolytica*. I'd like to thank Dr. Joseph Cheng and Dr. John Leavitt for their long discussions with me when I'm confused. And I'm especially grateful for Dr. James Wagner. It is impossible to complete my work without the tools he developed for *Y. lipolytica* and his knowledge. I'd also like to thank Dr. Kelly Markham, Claire Palmer and Lauren Cordova for providing

the TAL and ALA strains so I can complete the mating project. Thanks to everyone else for making Alper lab a fun place to do research.

I want to thank my two hard working undergraduate students, Olivia Schmitz and Sayuri Monarrez who worked with me. They made my life so much easier. I hope Olivia enjoy her life at Harvard Law School and wish Sayuri enjoy her new job.

I want to thank Dr. Dennis Wylie for the support on genome sequencing and Richard Salinas for assistance on flow cytometry.

I want to thank my previous supervisors, Dr. Ronnie Frederick, Dr. Katarzyna Gromek, Dr. Thomas Jeffries, Dr. Trey Sato and Dr. Yaoping Zhang. They trained me and prepped me to come to graduate school and become a scientist. They opened the gate of science for me.

Lastly, I want to thank my friends and families. I won't live such a happy life on foreign soil without all my friends. And it wouldn't be possible for me to be in this position without the support from my parents. They supported me to study what I love, and paid my tuitions anyway even they don't know what biochemistry and microbiology is about.

Abstract

Biochemical Production by *Yarrowia lipolytica* from Xylose

Haibo Li, PhD

The University of Texas at Austin, 2019

Supervisor: Hal Alper

Xylose is the second most abundant sugar in cellulosic hydrolysate. Xylose utilization has been extensively studied in both microbes that naturally utilize it, for example *Escherichia coli* and *Scheffersomyces stipitis*, as well as in model organisms that have abundant genetic tools such as *Saccharomyces cerevisiae*. However, in non-conventional microorganisms that possess great industrial potential, only minimal research has been conducted to enable and improve their xylose utilization. In this study, we chose *Yarrowia lipolytica* as an example to explore threekey strategies to enable xylose utilization by a non-catabolizing host for the purposes of producing a variety of value-added chemicals.

First, we integrated two heterologous genes that are essential for xylose utilization. Introduction of the pathway itself was not enough to enable xylose utilization, so a starvation-enabled adaptation was used to activate the pathway and kick-start xylose catabolism. Through whole genome sequencing, we found the copy numbers of both genes were multiplied during starvation. The resulting strain was able to produce 15g/L lipid from

xylose in bioreactors. We then tested several targets within the pentose phosphate pathway that have been shown to improve xylose utilization in *S. cerevisiae*. Among these, only native xylulose kinase overexpression was shown to effectively improve xylose utilization in *Y. lipolytica*. Next, we utilized this strain to expand the product portfolio from xylose through the use of a strain mating approach. Specifically, we altered the native mating type within this strain and mated it with three separately engineered strains of *Y. lipolytica* that overproduce α-linolenic acid, riboflavin or triacetic acid lactone. The mated diploid strains were able to produce all three products from xylose in high quantities. Finally, we turn to the first step of xylose catabolism, transport into the cell, to remove a common bottleneck. To do so, we utilized a series of evolution and selection steps to engineer a xylose transporter that is uninhibited by glucose when transporting xylose. This transporter can ultimately enable yeasts to co-utilize glucose and xylose.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and background

1.1 Lignocellulosic Biomass And Hydrolysate

Continued use of fossil fuels are unsustainable and human kind needs to find alternatives to this source of energy(Jones and Warner 2016). Bioenergy, along with solar energy, wind energy, hydropower and geothermal energy are the major renewable energies being explored to diversify our available energy sources(David Martin Alonso 2011; Dale and Ong 2012). However, fossil fuels are not only sources for energy, but are used as starting materials for plastic, synthetic rubber, chemical solvents, fabric, detergents, cosmetic products, pharmaceutical components and even food preservatives (Naik *et al.* 2010). In all of those products, fossil fuels provide physical material, which people are not able to obtain from sunlight, wind, flowing water, or heat from Earth (Tai 2012). However, biology can produce a diverse array of these physical substances from various carbon sources (Quinn and Davis 2015). Glucose is one of the most abundant sustainable carbon sources since it is the primary product from photosynthesis. Common glucose sources are sugar cane juice which is rich in glucose monomers, or hydrolysed corn or other crops that are high in starch. Both are currently commercially used for making bio-ethanol (Lynd 1996). Moreover, glucose from plant sources can provide a renewable source of carbon for bioproduction of a great number of compounds using engineered microorganisms (David Martin Alonso 2011). However, biofuels generated from crops, which are called first-generation biofuel, directly compete with food supplies(Naik *et al.* 2010). This competition has raised food prices in the international market in the past (McMichael 2010; Mol 2010). Thus, the search for cheap, non-competitive carbon sources led scientists to focus on non-edible lignocellulosic biomass.

Lignocellulosic biomass is dry plant matter used as the carbon source for second-generation biofuel (Naik *et al.* 2010; Brethauer and Studer 2015; Marriott, Gómez and Mcqueen-Mason 2016). The three main sources for lignocellulosic biomass are agriculture waste, energy crops, and wood (Lynd 1996). The most common agriculture wastes are corn stover, wheat straw, or rice straw (Donato *et al.* 2015). Energy crops are plants that can grow on marginal land that are not suitable for food crops, and produce high amount of biomass (Clifton-Brown *et al.* 2019). The most widely studied energy crop is switchgrass, which produces very large amount of biomass**.** Native prairie plants have also been studied for their environmental benefits (Boe and Lee 2007). The wood source for lignocellulosic biomass can be either hardwood, like poplar, or softwood, like pine trees (Przybysz Buzała *et al.* 2017).

On the structural/chemical level, lignocellulosic biomass is composed of three main components: cellulose, hemicellulose and lignin (Jin *et al.* 2013; Mcqueen-mason 2015; Marriott, Gómez and Mcqueen-Mason 2016). The specific ratio of these three components varies depend on the biomass source. All three components are polymers. Cellulose is polymer of glucose; hemicellulose is polymer of mainly pentose with hexose side chains and lignin is polymer of a group of aromatic compounds. Lignocellulose is generally pretreated prior to use to break down the polymeric material to monosaccharides (Chiaramonti *et al.* 2012; Zheng and Rehmann 2014). There are many types of pretreatment available, including physio-chemical pretreatment like ammonia fiber explosion (AFEX) (Balan *et al.* 2009), autohydrolysis (steam explosion) (Amiri and Karimi 2015), acid or alkali treatment (Takács *et al.* 2000; Iranmahboob, Nadim and Monemi 2002); chemical pretreatment like ionic liquids (Da Silva *et al.* 2013) or gamma-valerolactone (GVL) (Luterbacher *et al.* 2014); and enzymatic treatment with cellulase and hemicellulase (Zhang and Lynd 2006). After pretreatment, lignocellulosic biomass generally turns into a

dark, coffee-like liquid with contaminating material that could not be broken down. The residue is mostly lignin-derived compounds that unfortunately (at present) not effectively used to produce biochemicals at high level by human (Frei 2013). The dark liquid is called hydrolysate. Hydrolysate is a solution that contains high amount of sugars and also many other types of organic molecules including molecules that support cell growth and also many aromatic compounds that can inhibit microbial growth (Jin *et al.* 2013; Xue *et al.* 2018). The most abundant sugar in hydrolysate is glucose, and the second most abundant sugar is xylose(Gírio *et al.* 2010). This hydrolysate is used for microbial cell growth to produce the second-generation biofuel or other chemical molecules.

1.2 XYLOSE METABOLIC PATHWAY

Xylose ($C_5H_{10}O_5$) is a pentose sugar that makes up the polymer backbone for hemicellulose (Hilpmann *et al.* 2016). Within lignocellulosic hydrolysate, the average amount of xylose can reach to about 50% of the amount of glucose in hydrolysate (Gírio *et al.* 2010; Jagtap and Rao 2018). Additional sugars, for example, arabinose, mannose or galactose together make up a very small portion of the soluble sugar in hydrolysate (David Martin Alonso 2011; Jin *et al.* 2015a). Thus, if an industrial microorganism can use both glucose and xylose to make its product, the overall titer can increase up to 50% as compared to using glucose only. However, many of the microorganisms that are commonly used in industry including the main two either do not use xylose effectively (*E. coli*) (Ammar, Wang and Rao 2018) or not at all (*S. cerevisiae*) (Salusjärvi *et al.* 2008). This great potential leads many scientists around the world to study how to decode the ways for microbes to utilize xylose effectively.

Throughout the years, scientists have made significant progress on improving xylose utilization in microorganisms. The xylose utilization metabolic pathways have been

studied in microorganisms that naturally consume xylose and the enzymes catalyze these pathway reactions have been introduced into non-xylose utilizing model organisms like *S. cerevisiae*(Walfridsson *et al.* 1995; Anderlund, Rådström and Hahn-Hägerdal 2001; Jin and Alper 2005; Matsushika *et al.* 2009). There have been two main xylose pathways found and deeply studied, the oxido-reductase pathway and the isomerase pathway. The oxido-reductase pathway converts xylose into xylulose in two steps. First a xylose reductase (XR) converts xylose into xylitol, and then a xylitol dehydrogenase (XDH) converts xylitol into xylulose (Jeffries and Jin 2004). The isomerase pathway uses a single isomerase to directly convert xylose into xylulose, the common product for both pathways (Lee, Jellison and Alper 2012). Xylulose is then converted by action of a xylulose kinase (XKS) to xylulose-5-phosphate, which is incorporated into the native pentose phosphate pathway (Scalcinati *et al.* 2012). The microorganisms can then use their native genes to metabolize xylulose-5-phosphate.

1.3 *YARROWIA LIPOLYTICA* AS A CELL FACTORY

Yarrowia lipolytica is an oleaginous yeast emerging as one of the industry favorite non-conventional yeasts studied by both industrial and academic research groups (Blazeck *et al.* 2014; Gonçalves, Colen and Takahashi 2014; Liu *et al.* 2015a). Isolated from cheese (Roostita and Fleet 1996), *Y. lipolytica* has many desired properties for biochemical production (Palmer and Alper 2019). Like all oleaginous yeasts, *Y. lipolytica* is a natural high lipid producer. By applying the strategies of metabolic engineering, *Y. lipolytica* can produce nearly 99g/L lipid in bioreactor settings (Qiao *et al.* 2017). Lipids produced by this (and other) yeast can be transesterified with short chain alcohols to afford a mixture of compounds called biodiesel(Adesanya *et al.* 2014). Biodiesel can be used in diesel engines without making modifications on the engine (Abu-Hamdeh and Alnefaie 2015). Biodiesel

can also be used as aviation fuel (Taheripour *et al.* 2010). In contrast to bio-ethanol, which is easily outpaced by the use of batteries in electronic vehicles, biodiesel is more difficult to replace with alternative fuels for their use in aircrafts and railroads, and also their competitiveness in price (Carlson 2007). Besides lipids, *Y. lipolytica* has also been used for both protein and citric acid production (Gonçalves, Colen and Takahashi 2014). Recent research has also expanded the product portfolio to other oleochemicals like poly-unsaturated fatty acids or fatty alcohols (Cordova and Alper 2018), other organic acids like α-ketoglutaric acid, pyruvic acid, succinic acid and itaconic acid (Blazeck *et al.* 2015; Ryu, Hipp and Trinh 2015); other nutraceuticals like lycopene, β-carotene, (Matthäus *et al.* 2014; Czajka *et al.* 2018); and products derived from Acetyl CoA, like type III polyketides (Markham *et al.* 2018).

1.4 Enabling Xylose Utilization In *Y. lipolytica*

With such wide applications in chemical production, there remains one significant issue with *Y. lipolytica*: it does not naturally utilize xylose as a carbon source. The inability to use xylose limits the titer of products that can be produced from *Y. lipolytica*. Sequencing results have shown that a native XR and XDH are present in *Y. lipolytica* genome. However, *Y. lipolytica* is not able to natively grow in media in which xylose is the only carbon source (Ryu, Hipp and Trinh 2015). Some preliminary research has shown that a drop of xylose level can be detected after *Y. lipolytica* grows to high cell density when glucose was present, hinting that xylose utilization is possible in *Y. lipolytica*. Some early attempts to integrate xylose catabolic pathways into *Y. lipolytica*, however, did not result in a xylose utilization strain. The work detailed in Chapter 2 describes the successful construction of a xylose-utilizing *Y. lipolytica* strain. Xylose-catabolism was realized only after a starvation process which duplicated the copy numbers of xylose pathway genes that result in higher

expression levels. The xylose utilizing strain was able to make more than 15 g/L lipid when xylose was provided as the only carbon source in bioreactors. These results were published in *Biotechnology Journal* in 2016.

1.5 Improving Xylose Utilization Through Metabolic Engineering

After *Y. lipolytica* was made to be able to use xylose, we sought to improve this phenotype through additional metabolic engineering. Similar to *Y. lipolytica*, *S. cerevisiae* also cannot grow on xylose naturally, and after integration of xylose pathway genes into *S. cerevisiae*, it has been discovered that xylose utilization was not at maximum capacity due to inhibition, co-factor imbalance or metabolic bottle-necks (Sànchez Nogué and Karhumaa 2015). Scientists have identified a handful of targets that improve xylose utilization in *S. cerevisiae*, including xylulose kinase that is downstream of the integrated pathway and genes in pentose phosphate pathway (Walfridsson *et al.* 1995; Karhumaa, Hahn-Hägerdal and Gorwa-Grauslund 2005; Salusjärvi *et al.* 2008). The native versions of those genes were found in *Y. lipolytica* and tested in Chapter 3, along with additional genes that were hypothesized to be able to balance the co-factors in *Y. lipolytica* in xylose pathway. Our results demonstrate that native xylulose kinase overexpression most significantly improves xylose utilization in *Y. lipolytica*.

1.6 Expanding The Product Portfolio From Xylose In *Y. Lipolytica* Through Mating

One of the major advantages of *Y. lipolytica* is its diverse product portfolio. To achieve high product titers, many high production strains have undergone significant genetic modification, mutation and selection, or adaptation. Similarly, our xylose utilization strain also went through heterologous and homologous genes overexpression and starvation adaptation. Due to the limitation of genetic modification tools for *Y.*

lipolytica, transferring desirable phenotypes into a new strain background and maintaining effective expression levels presents a unique challenge (Liu *et al.* 2015a; Wagner and Alper 2015; Abdel-Mawgoud *et al.* 2018). To make products other than lipids from xylose, we used traditional mating methods to combine traits from different strains. *Y. lipolytica* have two mating types, type A and type B (Weber, Kurischko and Barth 1988; Barth 2011). *Y. lipolytica* is not able to switch mating type like *S. cerevisiae*, so we altered the xylose utilizing strain from mating type A into mating type B. The xylose strain after mating type switch was mated with three mating type A strains that produce drastically different products: α-linolenic acid, riboflavin and triacetic acid lactone. The mated diploid strain has high stability and produce very satisfying titer of products from xylose while maintaining titer from glucose. This work in Chapter 4 shows that mating can be a reliable method for expanding *Y. lipolytica* product portfolio.

1.7 Engineering Xylose Transporter Without Glucose Inhibition

One major issue for xylose utilization in microorganisms is that xylose utilization is inhibited by the presence of glucose (Subtil and Boles 2012; Farwick *et al.* 2014a). This phenotype has been observed in both *Y. lipolytica* and *S. cerevisiae* . Removing this glucose inhibition and allowing for glucose/xylose co-utilization can greatly increase fermentation efficiency. In Chapter 5, we demonstrated starting from an engineered xylose-only transporter, through multiple rounds of random mutagenesis, we were able to identify a few key residues that are related to glucose inhibition on xylose. In a model strain in which all other monosaccharide transporters were removed, this mutant transporter exhibited the capacity to transport both glucose and xylose with no inhibition to each other, but showed preference to xylose. This work was conducted in *S. cerevisiae*, and published in *Applied*

Microbiology and Biotechnology in 2016. This transporter can be further tested and applied in other xylose utilizing yeast strains.

1.8 SUMMARY

Due to the limited use of model yeast S. cerevisiae, scientist have drifting their focus away to other non-conventional yeasts for biochemical production. Among the non-conventional yeast, *Y. lipolytica* has drawn great attention due to its metabolic advantage and diversity in products making. In this work, we showed three complementary strategies involved in engineering *Y. lipolytica* into an effective xylose utilizing production host. The three steps include first introducing and activating xylose catabolic pathway; product portfolio expansion; and finding other targets to further improve xylose utilization. During the engineering process, we encountered many obstacles that are normally not seen in S. cerevisiae. To solve those problems, we used a combination of modern techniques like whole genome sequencing, site directed mutagenesis; traditional techniques like starvation induced adaptation, mating; as well as some new tools recently transplanted into *Y. lipolytica* system like piggyBac transposase to achieve our goal. In the end, we were able to obtains *Y. lipolytica* strains that are able to make 15 g/L lipid, 0.52 g/L α-linolenic acid, 96.6 mg/L riboflavin and 2.9 g/L triacetic acid lactone respectively as long as a transporter that are able to transport xylose and glucose simultaneously without inhibition to each other. This study can be used as a general guideline on how to enable a non-conventional yeast to utilize a non-native carbon source.

Chapter 2: Enabling xylose utilization in *Yarrowia lipolytica* for lipid production[1][1]

2.1 CHAPTER SUMMARY

The conversion of lignocellulosic sugars, in particular xylose, is important for sustainable fuels and chemicals production. While the oleaginous yeast *Yarrowia lipolytica* is a strong candidate for lipid production, it is currently unable to effectively utilize xylose. By introducing a heterologous oxidoreductase pathway and enabling starvation adaptation, we obtained a *Y. lipolytica* strain, E26 XUS, that can use xylose as a sole carbon source and produce over 15 g/L of lipid in bioreactor fermentations (29.3% of theoretical yield) with a maximal lipid productivity of 0.19 g/L/h. Genomic sequencing and genetic analysis pointed toward increases in genomic copy number of the pathway and resulting elevated expression levels as the causative mutations underlying this improved phenotype. More broadly, many regions of the genome were duplicated during starvation of *Yarrowia*. This strain can form the basis for further engineering to enhance xylose catabolic rates and conversion. Finally, this study also reveals the flexibility and dynamic nature of the *Y. lipolytica* genome, and the means at which starvation can be used to induce genomic duplications.

2.2 INTRODUCTION

An increased demand for energy and materials over the past few decades (Jones and Warner 2016) has been matched with depleting reserves of traditional fossil fuel, despite new mining technologies (Milano *et al.* 2016). To address this need, a significant

[1] Content in this chapter adapted from a previously authored publication written by HL. HL equally contributed to the contributed to experiments, analyses and writing of the manuscript.
Haibo Li and Hal S. Alper, 2016. Enabling xylose utilization in *Yarrowia lipolytica* for lipid production. *Biotechnology Journal* 11(9)1230-40.

amount of research has been placed on identifying alternative sources for liquid transportation fuels and chemicals, especially from lignocellulosic biomass (Lynd 1996; Naik *et al.* 2010; Brethauer and Studer 2015; Jin *et al.* 2015b; Mcqueen-mason 2015; Quinn and Davis 2015). Lignocellulosic hydrolysate (typically sourced from corn stover or switchgrass) is normally comprised of 60-70 g/L glucose and 30-40 g/L xylose (Pauly and Keegstra 2008). While the glucose portion can efficiently and rapidly be converted into fuel molecules by microbes, very few organisms are readily able to utilize xylose as effectively (Wang, Shopsis and Schneider 1980; Jeffries 1983; Jeffries and Jin 2004; Young, Lee and Alper 2010; Harner *et al.* 2015). As a result, metabolic engineering strategies for enabling microorganisms to use xylose has been the focus of research in the field (Jeffries and Jin 2004; Young, Lee and Alper 2010; Scalcinati *et al.* 2012; Parreiras *et al.* 2014; Harner *et al.* 2015; Ryu, Hipp and Trinh 2015; Sànchez Nogué and Karhumaa 2015).

One particular microbe of interest for biofuels production is the oleaginous yeast *Yarrowia lipolytica* due to its high capacity to convert glucose into lipids, a potential feedstock for biodiesel (Tai 2012; Blazeck *et al.* 2014; Gonçalves, Colen and Takahashi 2014). Additionally, *Y. lipolytica* has the potential to be a host system for a variety of other biochemical products (Nicaud *et al.* 2002; Blazeck *et al.* 2013a, 2015; Liu, Redden and Alper 2013). Previous results from the Alper lab have demonstrated that a combination of metabolic engineering and adaptive evolution can rewire *Y. lipolytica* to produce high titers (approaching 40 g/L) of lipids with cells containing nearly 90% of biomass as lipids (Liu *et al.* 2015b). However, the use of *Y. lipolytica* as a host cell

capable of utilizing lignocellulosic hydrolysate is complicated by the inability of this organism to effectively utilize xylose as a primary carbon source. Prior work has demonstrated engineering to enable some xylose utilizing capacity for this cell, however, the resulting strain did not grow well on xylose and the xylose utilization phenotype was not stable (Ryu, Hipp and Trinh 2015). Prior results from our laboratory indicate that xylose can be converted into lipids only when cells are transferred in stationary phase into high concentrations of xylose (Blazeck *et al.* 2014). Even still, this cell is unable to use xylose as a primary source for conversion to biomass. Additional literature and research findings also corroborates the findings that xylose catabolism is not fastidious in this organism. Thus, pathway engineering is necessary to improve the xylose catabolic phenotype in *Y. lipolytica*.

Xylose catabolism can be achieved by two major pathways: the oxidoreductase pathway and the isomerase pathway (Young, Lee and Alper 2010). Both of these pathways have been well-described and documented previously (Jeffries 1983; Jeffries and Jin 2004). Briefly, the oxidoreductase pathway (most studied from *Scheffersomyces stipitis*) consists of two major steps to first convert xylose into xylitol by xylose reductase (*XYL1*), and then xylitol into xylulose by xylitol dehydrogenase (*XYL2*). In the isomerase pathway (initially identified in *Lactobacillus pentosus* and studied from *Escherichia coli* and *Piromyces sp.*) consists of one step that reversibly converts xylose into xylulose by xylose isomerase (*XylA*). In both pathways, the resulting xylulose is then phosphorylated by xylulokinase (*XYL3)* into xylulose-5-phosphate before entering the pentose phosphate pathway. Among these two pathways, *Y. lipolytica* has putative genes for *XYL1*, *XYL2*

and *XYL3*, but only *XYL3* has been proven to be functional and expressed (Ryu, Hipp and Trinh 2015). In a previous report, activation of the endogenous *XYL1* and *XYL2* genes in *Y. lipolytica* was possible, albeit with a poor resulting growth rate and strain stability (Ryu, Hipp and Trinh 2015). As a result, it is likely that heterologous pathways would enable *Y. lipolytica* to utilize xylose stably and efficiently.

In this study, we evaluate both potential pathways for xylose catabolism in *Y. lipolytica* and demonstrate that codon-optimized versions of *S. stipitis XYL1* and *XYL2* can function in *Y. lipolytica*. In contrast, a codon-optimized version of the isomerase pathway was not found to be active. Next, we demonstrate that an adaptive, evolutionary engineering period (induced by starvation) is necessary to generate cells that can optimally use xylose in a stable manner. Through genome resequencing, we find that this efficient phenotype was predominately enabled by gene duplications in the *XYL1* and *XYL2* to allow for higher expression levels. By performing this optimization in a previously engineered lipid overproducing strain (E26), we demonstrate the capacity of the resulting cells to produce 15 g/L of lipids with productivities of 0.19 g/L/h in bioreactor fermentations of xylose. This represents a nearly 30% conversion yield of xylose and demonstrates the potential to enable xylose catabolism in *Y. lipolytica* for fuels and chemicals production.

2.3 RESULTS

Initial benchmarking growth assays were performed with strains and in contrast to previous reports (15, 32, 33), we found that *Y. lipolytica* strains PO1f and E26 were unable to grow in minimal media with xylose as the only carbon source. This result is consistent

with additional reports in literature (Ruiz-Herrera and Sentandreu 2002; Pan *et al.* 2009). Our previous results demonstrating xylose conversion in stationary phase cells (Blazeck *et al.* 2014) suggests an unknown regulation and temporal activation of the endogenous oxidoreductase pathway. To avoid any of these complications, we opted to utilize heterologous pathway expression of both the oxidoreductase pathway (codon-optimized *XYL1* and *XYL2* from *S. stipitis*) and the isomerase pathway (codon-optimized *XylA* from *Piromyces sp.*). Initial tests were conducted to determine the effectiveness of these pathways.

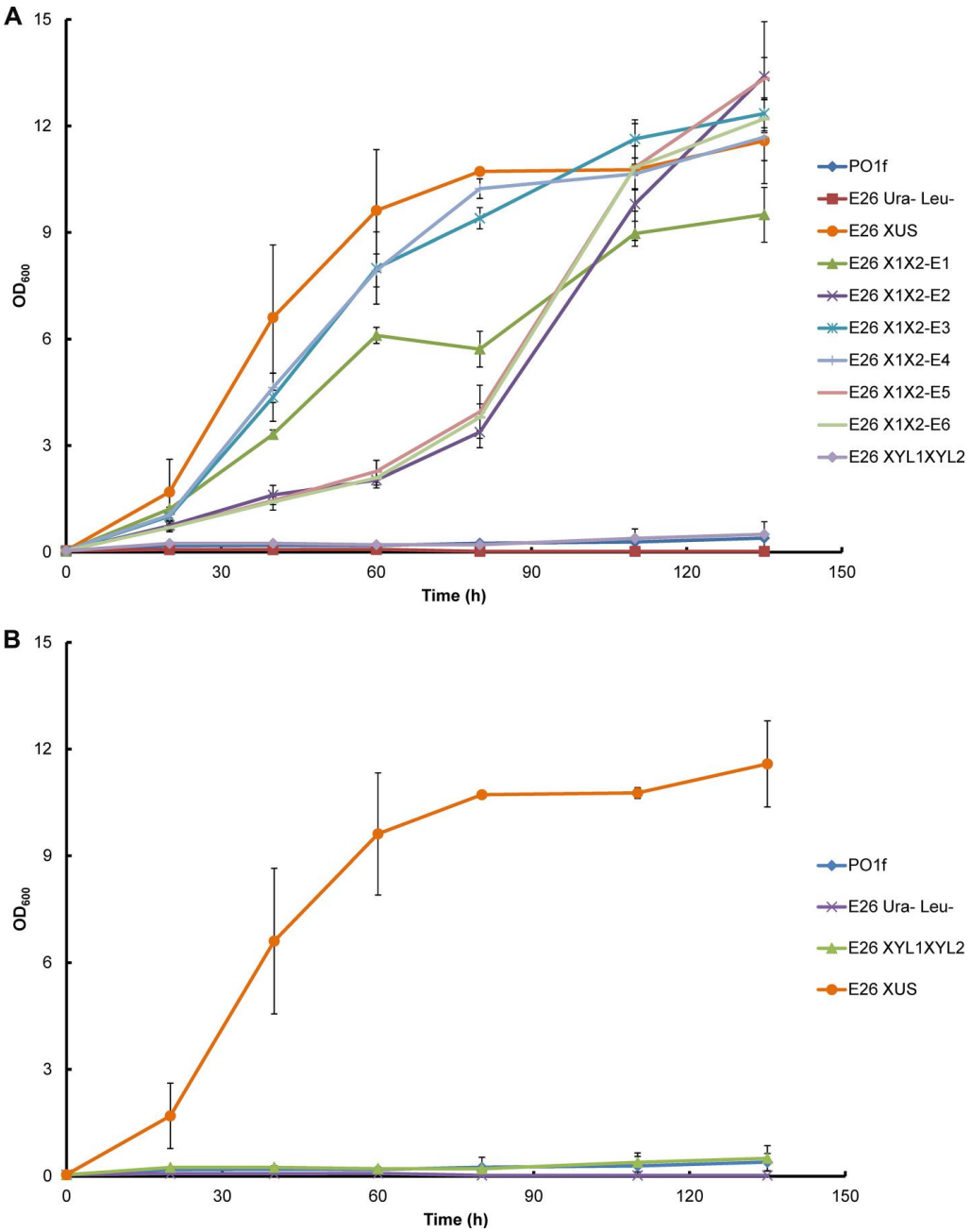
2.3.1 Heterologous oxidoreductase pathway complementation and starvation enables xylose catabolism

Heterologous xylose pathways were introduced into both the wild-type (PO1f) and engineered (E26) strains of *Y. lipolytica*. Complementation of either pathway (oxidoreductase or isomerase) was insufficient to enable growth of any strain of *Y. lipolytica*. Thus, we sought to perform a starvation-enabled adaptive evolution. Specifically, each of these strains were grown in glucose and allowed to sit in carbon-depleted medium for 14 days. After this time, starved cells were washed with water, and then transferred into fresh 50 mL cultures of YNB with 20 g/L xylose. After 12 days of culture, strains complemented with the oxidoreductase pathway were able to consume all of the xylose present. This phenomenon occurred with both the wild-type (PO1f) and engineered (E26) strains. In contrast, this utilization was not seen with the isomerase strains as starved cells showed no difference in OD after incubated in xylose for 12 days

and no xylose consumption could be measured. This indicates that codon-optimized isomerase does not work in *Y. lipolytica* without further engineering.

Next, individual colonies adapted from the oxidoreductase pathway were isolated and screened for their capacity to grow on xylose. In particular, seven individual colonies isolated from the adapted engineered strain were confirmed to have strong xylose catabolic features and thus were further characterized in shake flasks (Figure 2-1). The diverse growth patterns observed among these strains suggests that they may not be genetically identical and thus may represent unique evolutionary trajectories. The best performing strain was renamed E26 XUS (the E26 Xylose Utilizing Strain) and was selected for further characterization including fermentation capacity and genome resequencing.

Figure 2-1: Growth profiles of isolated *Y. lipolytica* strains on 20 g/L xylose.



(A) Growth curves of all seven isolated E26 derived strains expressing XYL1 and XYL2 after starvation are shown compared with PO1f, E26 Ura- Leu- and E26 XYL1XYL2 as controls. (B) The full growth curve of the E26 XUS strain is shown with the same controls as in (A). Each data point represents an average $\pm$ 1 standard deviation of three biological replicate experiments that were grown in parallel.

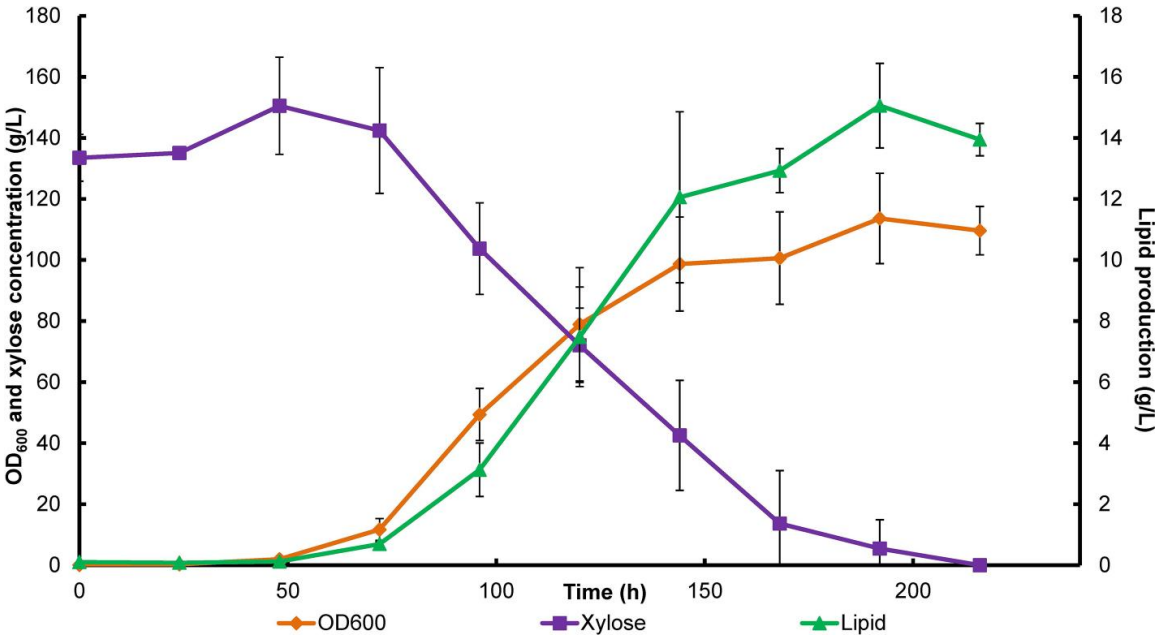
2.3.2 E26 XUS enables high lipid production in bioreactors

Following isolation, we sought to evaluate the lipid production traits of the E26 XUS strain in a bioreactor setting. For this experiment, we selected the controlled pH, temperature and dissolved oxygen settings previously found optimal for lipid production using glucose as carbon source (Blazeck *et al.* 2014). Moreover, to best compare the resulting performance to the parental strain on glucose, we used the same media formulation swapping glucose for xylose. Specifically, this media contained 160 g/L xylose, 10 g/L ammonium sulfate, and YNB (Liu *et al.* 2015b). The resulting fermentation was able to reach an OD_{600} of 100 within 120 hours and all xylose was consumed before 216 hours (Figure 2-2). While these cultures seemed to lag by nearly 24 hours compared to growth on glucose, these results demonstrate the strongest growth of *Y. lipolytica* on xylose reported to date. Total lipid production of just over 15 g/L was achieved (Table 2-1) with a maximal lipid productivity of 0.19 g/L/h realized. A lipid content of mainly C_{18} and C_{16} fatty acid species were found when converting xylose into lipids (Table 2-1) which is similar to the composition seen with glucose conversion (Blazeck *et al.* 2014). Collectively, this strain was able to achieve 29.3% of the theoretical yield of xylose into lipids. While these overall metrics lag those achieved in glucose, they represent a strong initial xylose catabolic capacity that has not been previously reported.

Table 2-1: Lipid production data from last time point of bioreactor fermentation.

16:0	16:1	18:0	18:1	18:2	Total (g/L)
3.49±0.31	0.53±0.14	3.52±0.47	6.26±0.76	1.25±0.16	15.06±1.60

Figure 2-2: Bioreactor fermentation of the E26 XUS strain in xylose.



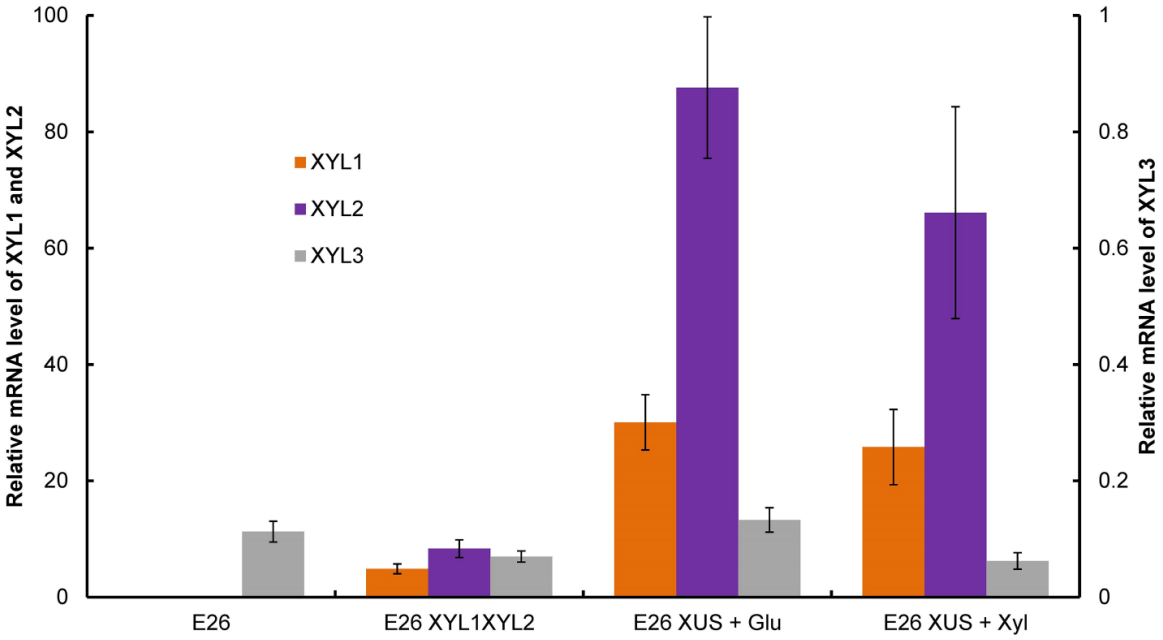
Growth curve, xylose utilization data and lipid production data in 1.5 L bioreactor fermentations. Each data point represents an average ± 1 standard deviation of three individual bioreactor fermentations.

2.3.3 *XYL1* and *XYL2* are highly expressed in the XUS strain

Next, we performed a series of genetic analysis in an effort to identify the causative genetic factors leading to the xylose catabolic phenotype observed here. First, we performed focused Sanger sequencing of the *XYL1* and *XYL2* genes present in PO1f *XYL1XYL2*, PO1f XUS, E26 *XYL1XYL2*, and E26 XUS strains. Surprisingly, no mutations were found in any of these strains, indicating that the xylose catabolic phenotype was not due to mutations in the enzymes.

After pathway specific mutations were ruled out, we next evaluated the expression level of these enzymes across the strains and conditions. Previous work suggests that endogenous, putative *XYL1* and *XYL2* can be upregulated in xylose utilizing strains (Ryu, Hipp and Trinh 2015). Moreover, an additional study suggests that heterologous expression of xylose pathway enzymes can have higher expression than endogenous *XYL1* and *XYL2* (Tai 2012). To evaluate expression of our heterologous enzymes, we conducted RT-PCR. Specifically, we observed that the expression level of exogenous *XYL1* and *XYL2* in starvation-treated E26 XUS strains is indeed much higher than in the E26 *XYL1XYL2* strain before starvation (Figure 2-3). We also measured the *XYL3* expression level under the same condition, and observed no substantial changes of *XYL3* expression between the strains (Figure 2-3). The expression of *XYL1 and XYL2* was consistent regardless of whether the cells were grown in glucose or xylose, a trait that is expected due to the constitutive TEF-based promoter used to drive expression. Thus, these results indicate that expression of *XYL1* and *XYL2* increased 6-fold and more than 10-fold, respectively in the evolution of the E26 XUS strain.

Figure 2-3: Expression levels of *XYL1*, *XYL2* and *XYL3*.



Relative mRNA expression levels of *XYL1*, *XYL2* and *XYL3* at exponential phase in 50 mL of CSM in 125 mL flask for strain E26, E26 *XYL1XYL2* and E26 XUS were measured with RT-PCR. E26 and E26 *XYL1XYL2* were grown on glucose. E26 XUS was grown on either glucose or xylose. Actin DNA copy number was used as internal reference. The expression level of *XYL1* and *XYL2* is shown with the primary axis on the left, and expression level of *XYL3* is shown with secondary axis on the right. Each data point represents an average value ± 1 standard deviation calculated from the results of three biological replicates.

2.3.4 Genome resequencing reveals *XYL1* and *XYL2* duplication as a major factor for improved xylose consumption

To further understand the underlying genetic changes responsible for improved xylose utilization, we re-sequenced the entire genome of the E26 XUS strain and compared this sequence to the E26 sequence we described previously (Liu *et al.* 2015a, 2015b). After SNP analysis, we found that only one absolute mutation (S713L) occurred in an annotated open reading frame (in YALI0F01562p). YALI0F01562p is a putative protein on the F chromosome that shows similarity to some C6 transcription factors and nitrogen regulators by BLAST homology search (Madden 2003) . Interesting, we did not see this same mutation based on Sanger sequencing in the other alternative xylose utilizing strains isolated in this work (strains shown in Figure 2-1). Thus, while potentially interesting, this mutation was not a ubiquitously dominant, causative mutation for improving xylose utilization.

Beyond SNPs in the genome, we also observed that many regions of the genome have significantly higher reads relative to their neighboring regions. Moreover, some of these regions had mutations in only a fraction of the reads indicating the presence of duplication events / repeated elements with the genome during this process. Specifically, we found 28 regions that were duplicated with mutations in their duplicated copies, and hundreds with duplications but no mutations. The sizes of the duplicated pieces ranged from 100 base pairs to several kilobases. Due to the inability to fully scaffold most of these repeated reads due to their length, it is not possible to exclude the possibility that some of the repeated pieces of DNA were inserted into regions related to xylose utilization.

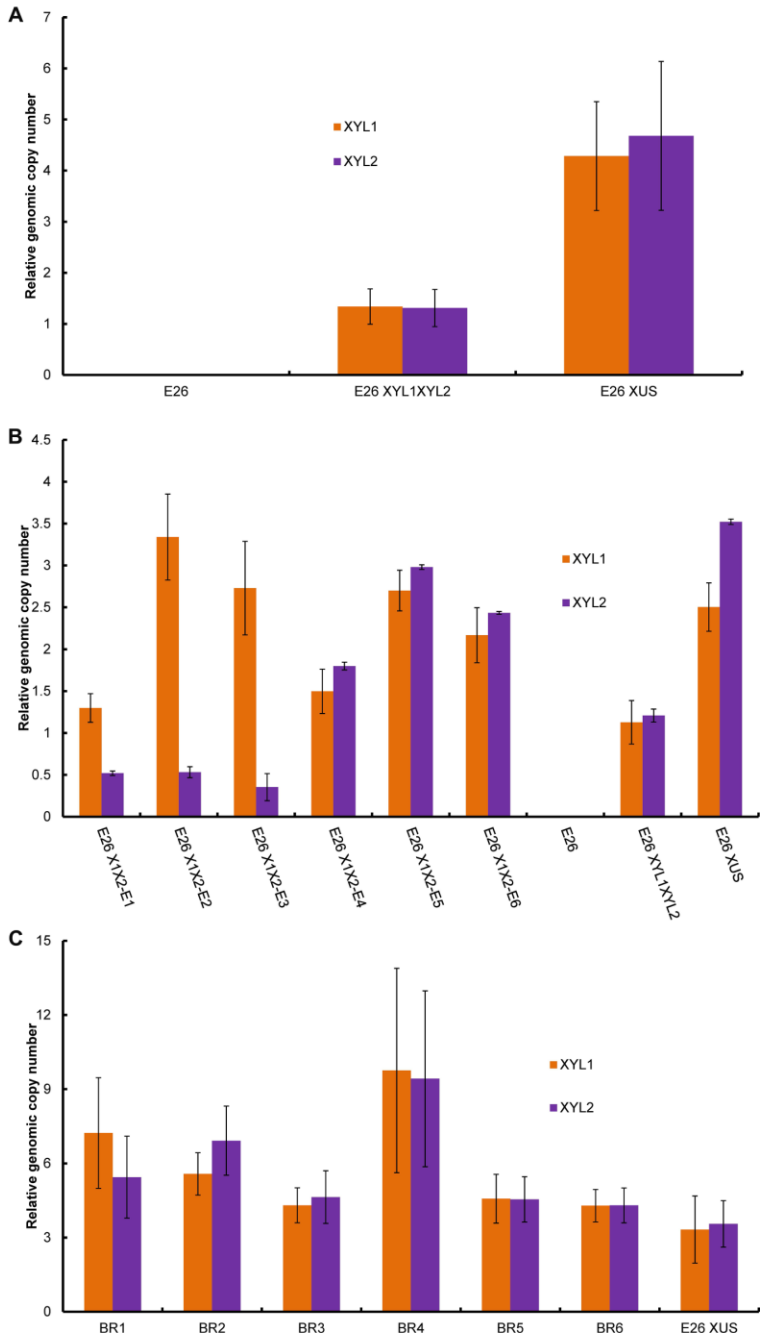
Furthermore, it became clear that multiple copies and reads of the *XYL1* and *XYL2* genes were present. This observation together with the lack of any mutations in the promoter regions of these genes would lead to the increase in expression seen in Figure 2-3.

To validate an increased copy number of *XYL1* and *XYL2*, we performed a qPCR assay on genomic DNA for copy number of the *XYL1* and *XYL2* genes before and after the starvation adaptation. Indeed, we found that the copy number of *XYL1* in E26 XUS was 3.2 fold higher than in E26 *XYL1XYL2*, and the copy number of *XYL2* in E26 XUS was 3.6 fold higher than in E26 *XYL1XYL2* strain (Figure 2-4A). To further confirm that this observation of increased copy number of genes was genuine across all isolated strains, we measured the copy number of *XYL1* and *XYL2* in the other 6 strains that can grow on xylose and described in Figure 2-1. From this analysis, we found that most of these strains have at least one gene with more copies than the parental strain (Figure 2-4B). Furthermore, there was a strong correlation between the actual growth on xylose (Figure 2-1) and increased copy numbers for both *XYL1* and *XYL2* (Figure 2-4B).

As a result, the genome sequencing and copy number assays demonstrates an increase in copies of the *XYL1* and *XYL2* genes that may have been generated through the same mechanism as the other repeat elements found within the genome. This observation indicates that the genome of *Y. lipolytica* may in fact be a rapidly evolving and potentially dynamic system at times. As evidence for this fact, we conjectured that it may be possible that during the 216 hour bioreactor fermentation, further genome restructuring may occur due to the selection pressure on xylose. Moreover, we did observe some bioreactor runs that were more efficient than others leading to this suspicion. To confirm this finding, we

performed the copy number assay on *XYL1* and *XYL2* from genomic DNA extracted at the end of the bioreactor fermentations from six independent trials. This analysis did show that copy number can vary by up to 3-fold across trials. Moreover, we saw a distribution of both stable genotypes (copy numbers of nearly 3 as seen in the tested E26 XUS strain) and further enriching genotypes (further duplicating genotypes leading to higher copy number) (Figure 2-4C). Collectively, these results point both toward the malleability of the *Yarrowia* genome as well as the potential for further improvements to this strain.

Figure 2-4: Relative genomic copy numbers of *XYL1* and *XYL2*.

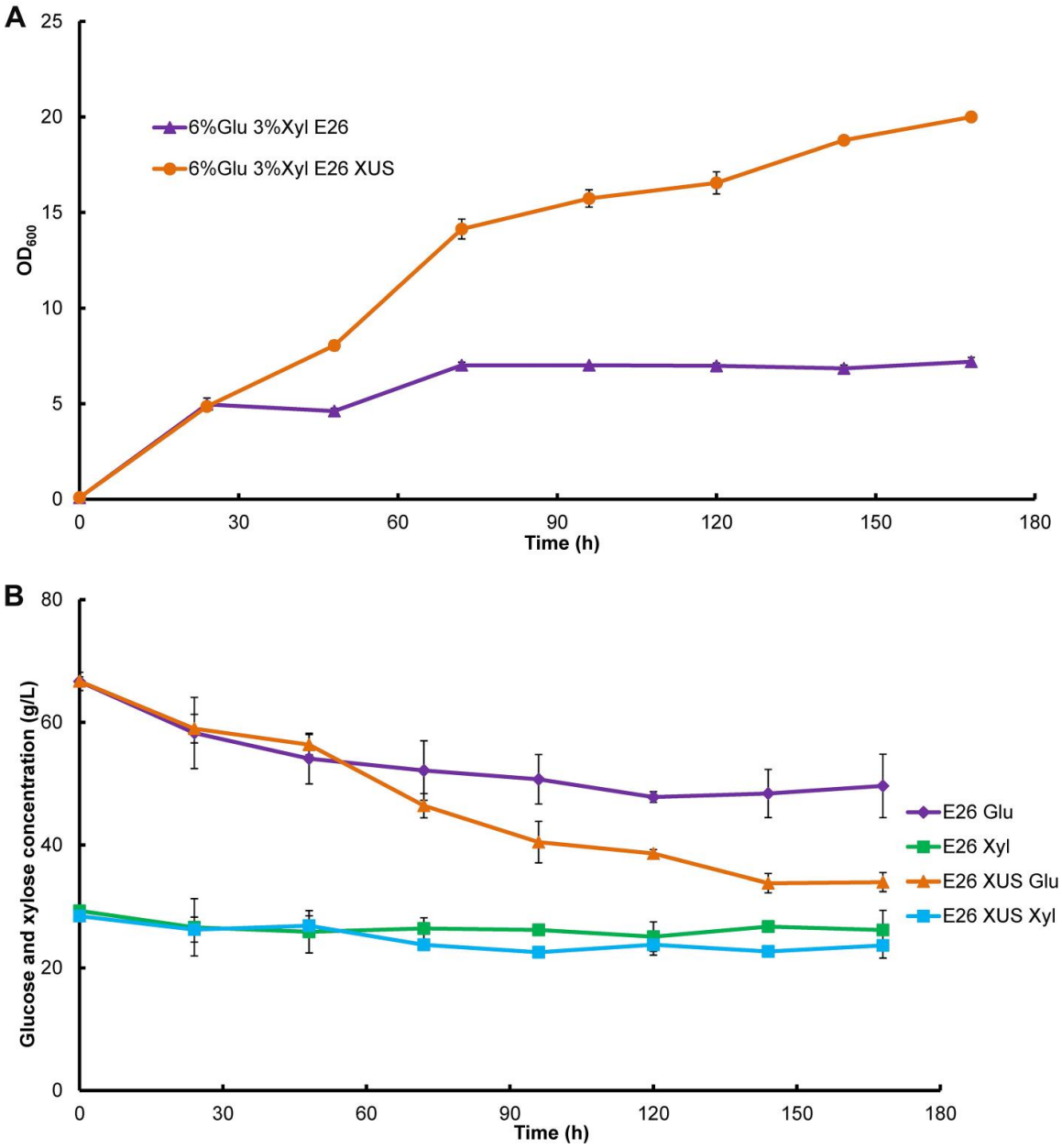


(**A**) Relative*XYL1* and *XYL2* copy number in strain E26 XUS. (**B**) Relative *XYL1* and *XYL2* copy number in different isolated E26 derived strains as shown in Figure 2-1. (**C**) Relative *XYL1* and *XYL2* copy number from 192 h samples of different bioreactor fermentations are displayed. Genomic copy number of actin gene was used as reference. Each data point represents an average value ± 1 standard deviation calculated from the results of three biological replicates (**A**) or three technical replicates (**B** and **C**).

2.3.5 Evaluation of E26 XUS for xylose-glucose co-utilization fermentation

After having characterized the E26 XUS strain and identifying the underlying genetic factors, we sought to further evaluate its potential in glucose-xylose co-utilization, a significant barrier in bio-ethanol production using genetic modified *S. cerevisiae* (Subtil and Boles 2012). Prior studies have claimed that *Y. lipolytica* xylose utilizing strains are able to co-utilize glucose and xylose at low concentrations (Ryu, Hipp and Trinh 2015). Previous work on sugar transporters from *Y. lipolytica* in our laboratory suggest that they are more specific for glucose than xylose when heterologously expressed (Young *et al.* 2011, 2014). To thoroughly test co-utilization properties, we tested our E26 XUS strain in YNB + CSM media with 60 g/L glucose and 30 g/L xylose, a concentration and ratio that best mimics the hydrolysate produced from corn stover (Dale and Ong 2012). In contrast to previously reported studies, we did not observe any appreciable xylose utilization in the presence of 60 g/L glucose (Figure 2-5). These results align with our prior transporter studies suggesting that xylose utilization in *Y. lipolytica* is inhibited by glucose, at a minimum, at least at high concentrations. However, it was noted that the E26 XUS strain grew significantly better than E26 strain in the mixed sugar media with a marked improvement in glucose utilization rate. This observation is indeed promising and while flask experiment may not directly reflect behavior in bioreactors, it is worth investigating the potential of this strain in glucose in a follow-up experiment.

Figure 2-5: Mixed sugar utilization test on E26 XUS.



(**A**) Growth profile of E26 vs. E26 XUS on 60 g/L glucose and 30 g/L xylose. (**B**) Glucose and xylose concentration in the mixed sugar fermentations were measured. Each data point represents an average value ± 1 standard deviation calculated from the results of three biological replicates.

2.4 DISCUSSION

Work within in this chapter, we established a strain of *Y. lipolytica* that is able to utilize xylose as a sole carbon source for growth, and convert this sugar into intracellular lipids. This result was obtained by first introducing heterologous *XYL1* and *XYL2* genes from *S. stipitis* into *Y. lipolytica* and second enabling gene duplication that occurred through starvation. The resulting strain, E26 XUS, was able to grow to OD_{600} above 100 in bioreactor fermentations and produced nearly 15 g/L of lipid at a maximum productivity of 0.19 g/L/h. While this strain produced less lipid than observed from similar concentrations of glucose (39.1 g/L from glucose (Liu *et al.* 2015b)), the strain was able to support higher growth in glucose/xylose mixed sugar media relative to the parental E26, suggesting further interesting features with the genome rearrangements.

Throughout this experiment, it was clear that increases to the copy number of this heterologous pathway led to increased expression and improved phenotype. However, it is important to note that the observed expression level of *XYL1* and *XYL2* did not directly correspond (i.e. 1:1 correlation) to the copy number increase, thus integration loci plays an important role. The integration plasmids used in this study contained 26S ribosomal DNA regions flanking the expression cassette. At first, it was presumed that multiple 26S sites in the *Y. lipolytica* genome were the likely source of these multiple integrations. However, TAIL-PCR was unable to confirm that the initial integrations nor the later duplicated ones have inserted into 26S regions. In contrast, we did find evidence of a potential hot spot for genomic integration in *Y. lipolytica*. Specifically, in PO1f XUS strain, at least one copy of *XYL1* has integrated into Chromosome C before nucleotide

45962, and at least one copy of *XYL2* of E26 XUS was also integrated into this same locus. It should be noted that none of the original copies in the *XYL1XYL2* strain were integrated into this spot suggesting that this location may be a hot spot for duplication activity. Regardless, this work further provides yet further proof that the efficiency of homologous recombination in *Y. lipolytica* is very low and that short flanking pieces of DNA do not result in specific, targeted integration in *Y. lipolytica*.

The very existence of multiple duplications (even evolving ones throughout the course of a fermentation) suggests that the genome of *Y. lipolytica* is quite dynamic at times. The observation that starvation of the cell led to high numbers of duplication regions (28 with mutations and hundreds without) along with sizes of upward of several kilobases demonstrates that the duplication event is not specific to heterologous DNA. As the exact mechanism of the duplication and precise locations of the duplicated pieces are still unknown, and it is not clear how vastly altered the overall chromosomal structure of *Y. lipolytica* can become. Nevertheless, this high adaptability of the *Y. lipolytica* genome can be a great advantage when doing metabolic engineering and strain improvement. At the same time, it could also present challenges for stable industrial production.

The creation of a xylose utilization phenotype through starvation was possible in both the engineered E26 and wild-type PO1f strains. Thus, it can be concluded that the observed duplications were not a product of the engineering and EMS mutagenesis used to construct the E26 strain. As a result, prior attempts at creating xylose utilizing strains using 4 and 10 day “adaptation” steps (Tai 2012; Ryu, Hipp and Trinh 2015) may have likely led to similar duplications.

In this work, we were not able to construct a xylose utilizing strain using the xylose isomerase pathway despite success with this enzyme in *S. cerevisiae* (Lee, Jellison and Alper 2012, 2014). It is unclear what limitations exist for making this pathway functional and this enzyme may require mutations for stability. Future work that can activate this pathway would take advantage of the co-factor free nature of the isomerase pathway compared to the oxidoreductase pathway (Dulermo *et al.* 2015).

Finally, the observed growth kinetics and sugar utilization rates indicated that the xylose utilization in the E26 XUS strain was strongly inhibited with conditions of 60g/L glucose and 30g/L xylose. As a result, further work is necessary to remove the inhibition by glucose to enable a co-fermentation. Nevertheless, the resulting titer and yield of lipid production from xylose achieved here are appreciable with future promise for improved productivity through additional strain engineering.

Chapter 3: Improving Xylose Utilization in *Yarrowia lipolytica* through Metabolic Engineering

3.1 INTRODUCTION

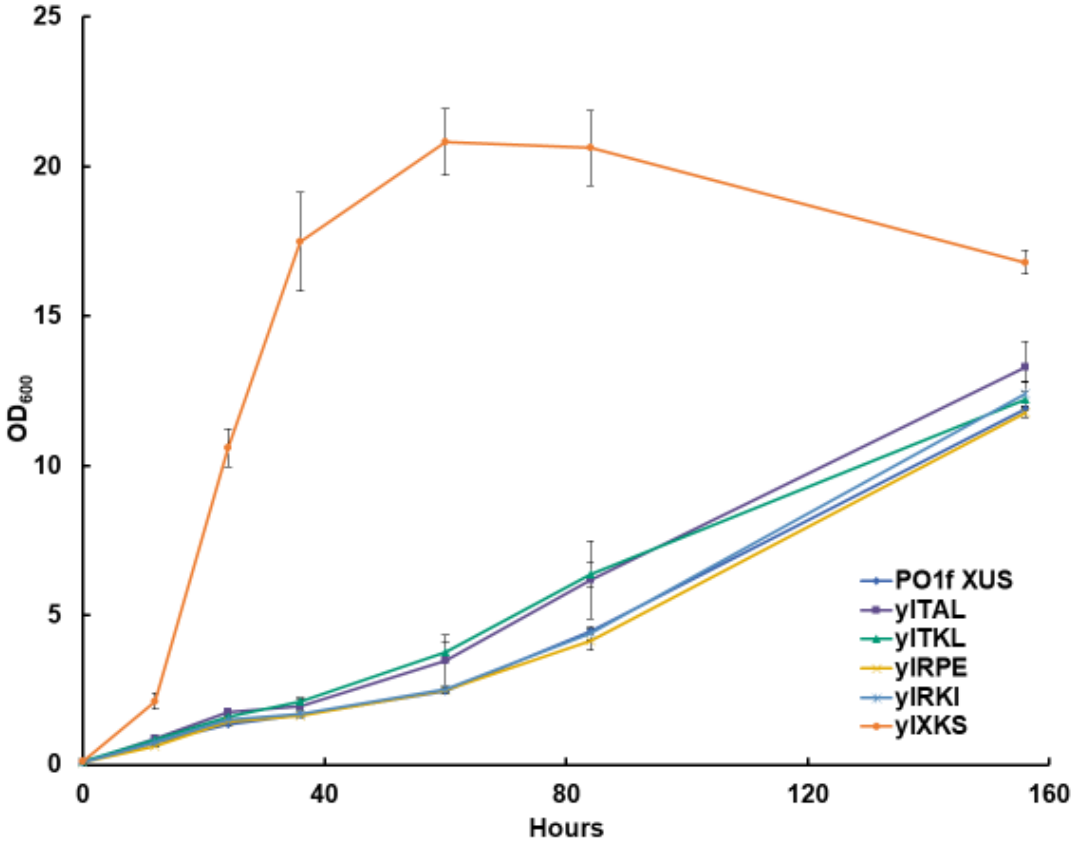
Improving the conversion of xylose by the unconventional yeast *Yarrowia lipolytica* has the potential to improve the economics of product production. To this end, researchers have attempted to improve sugar utilizing efficiency through metabolic engineering. Enabling *Y. lipolytica* to utilize xylose, the second most abundant sugar in biomass hydrolysate, can theoretically improve titer as much as 50% on total biomass (Pauly and Keegstra 2008). Results discussed in Chapter 2 demonstrated the ability to enable *Y. lipolytica* to utilize xylose and produce lipids from this xylose (Li and Alper 2016). In this chapter, several different pathway engineering applications are undertaken to further improve *Y. lipolytica* growth on xylose. First, several genes downstream of XYL1 and XYL2 were overexpressed to increase carbon flux downstream. These results uncover that overexpressing native XKS drastically improved the growth in xylose whereas overexpressing the other four genes in pentose phosphate pathway (PPP) did not make a difference. Next, attempts were made to improve growth in xylose by balancing cofactors. In previous fermentation runs, xylitol was identified in the end product, which indicates that not all of the xylose was utilized fully due to a co-factor imbalance. Two approaches were tried to balance the cofactors. However, introducing NADH phosphatase or replacing NADPH preferred XYL1 with an NADH preferred XYL1.2 from *Spathaspora passalidarum* both did not improve growth on xylose.

3.2 RESULTS

3.2.1 Improving xylose utilization in *Y. lipolytica* by overexpression of ylXKS and pentose phosphate pathway targets

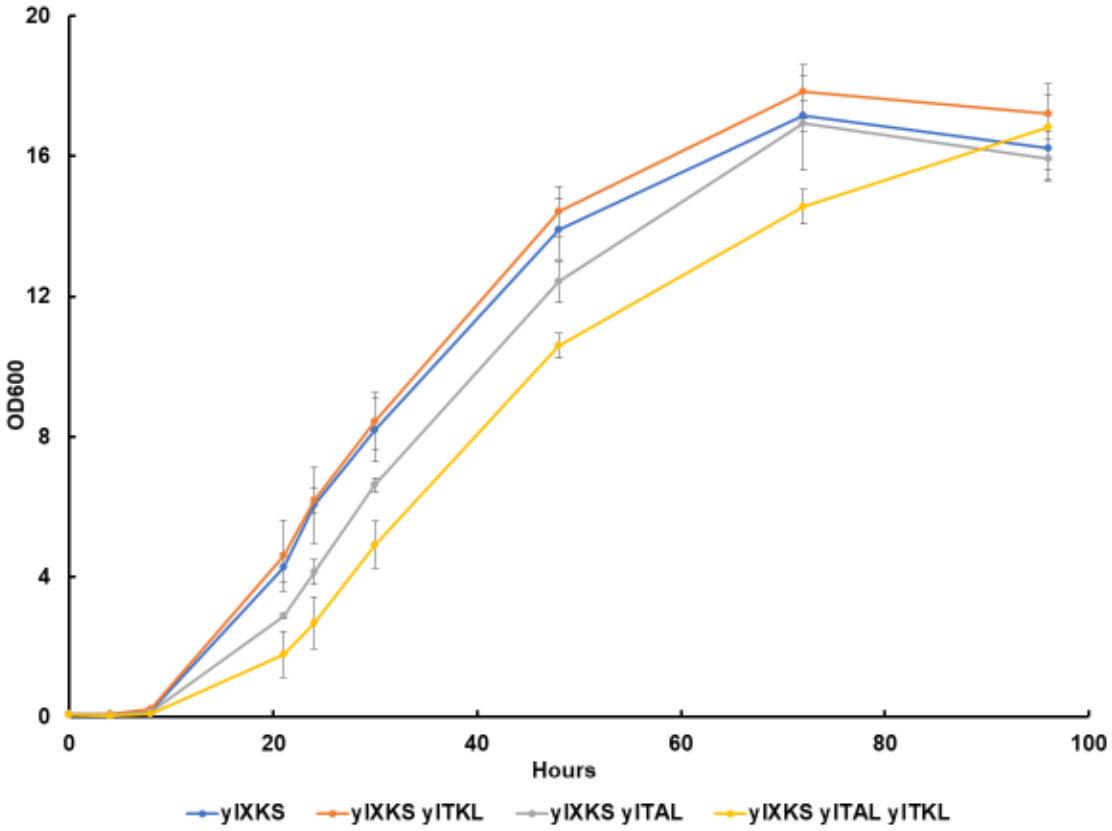
In prior work, we established a strain of *Y. lipolytica* capable of converting xylose into lipids through the integration of a xylose reductase and xylitol dehydrogenase along with a starvation adaptation (Li and Alper 2016). While these strains were suitable to utilize xylose, they still suffered from slow conversion. To improve catabolic rates, we explored the overexpression of the native xylulose kinase (ylXKS: YALI1F14583g) and four native pentose phosphate pathway related genes (ylTAL: YALI1F20914g, ylTKL: YALI1E07744g, ylRPE: YALI1C16621g and ylRKI: YALI1B09025g). Constructs overexpressing these genes through the use of a strong hybrid promoter (Blazeck *et al.* 2011) were integrated into an evolved xylose utilizing strain, PO1f XUS. In an analogous manner to *S. cerevisiae*, ylXKS overexpression greatly improved both growth rate and final OD in the presence of xylose (Figure 3-1). This strain reached an OD600 of more than 20 in 60 hours whereas the PO1f XUS strain only reached an OD of less than 12 at the 156 hour timepoint. The overexpression of ylTAL and ylTKL both individually exhibited a slight, but significant, improvement in growth on xylose whereas overexpression of ylRKI and ylRPE did not improve growth rate or level. An experiment was conducted to stack these targets; however, the results were not significantly distinct from just ylXKS alone (Figure 3-2).

Figure 3-1: Cell growth of PO1f XUS strain overexpressing genes from pentose phosphate pathway



Growth curves of PO1f XUS strain overexpressing native transaldolase (TAL), transketolase (TKL), ribose 5-phosphate ketol-isomerase (RKI), ribulose 5-phosphate epimerase (RPE) and xylulokinase (XKS) are shown compared with PO1f XUS as control. Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

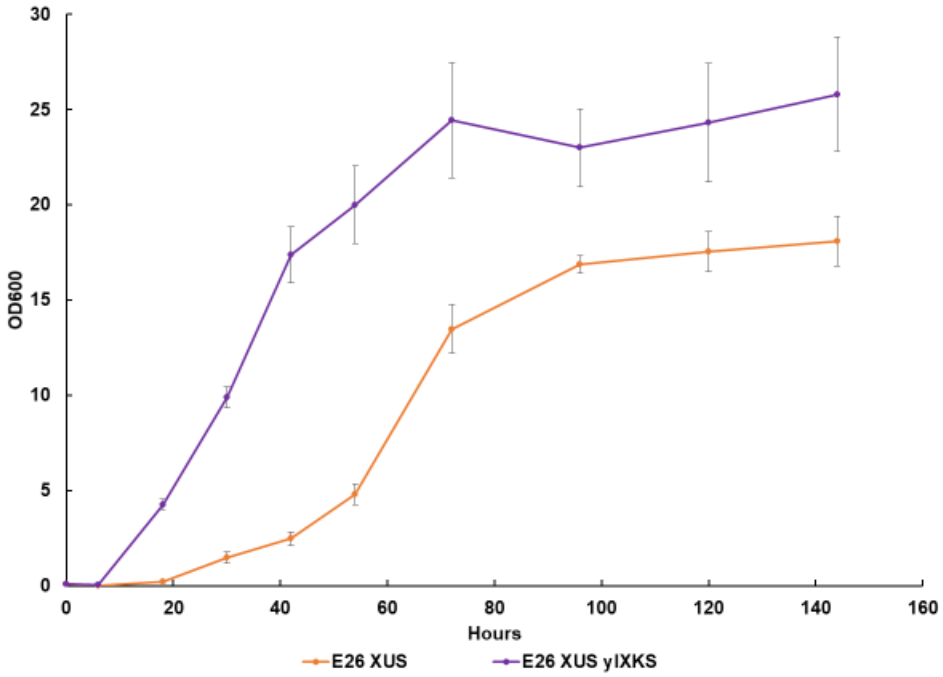
Figure 3-2: Cell growth of PO1f XUS overexpressing native transaldolase (TAL), transketolase (TKL), or both in addition to overexpressing XKS compared to PO1f XUS strain overexpress native XKS only.



Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

Next, we tested the overexpression of ylXKS in a high lipid producing, xylose utilizing strain (E26 XUS), and saw a similar trend with greatly improved growth on xylose (Figure 3-3). As a result, we then chose this particular high lipid xylose strain for mating type swap in Chapter 4. The overall function and importance of ylXKS in *Y. lipolytica* has been showed in the literature (Rodriguez *et al.* 2016). The study conducted here further confirms that ylXKS serves as a rate-limiting bottle neck for xylose growth. Moreover, the limited impact of other pentose phosphate pathway target overexpressions is corroborated by the native propensity for this strain to divert part of its carbon flux into the PPP for NADPH generation during lipid production (Wasylenko, Ahn and Stephanopoulos 2015a)..

Figure 3-3: Cell growth of E26 XUS strain overexpress native XKS compared with E26 XUS as control.

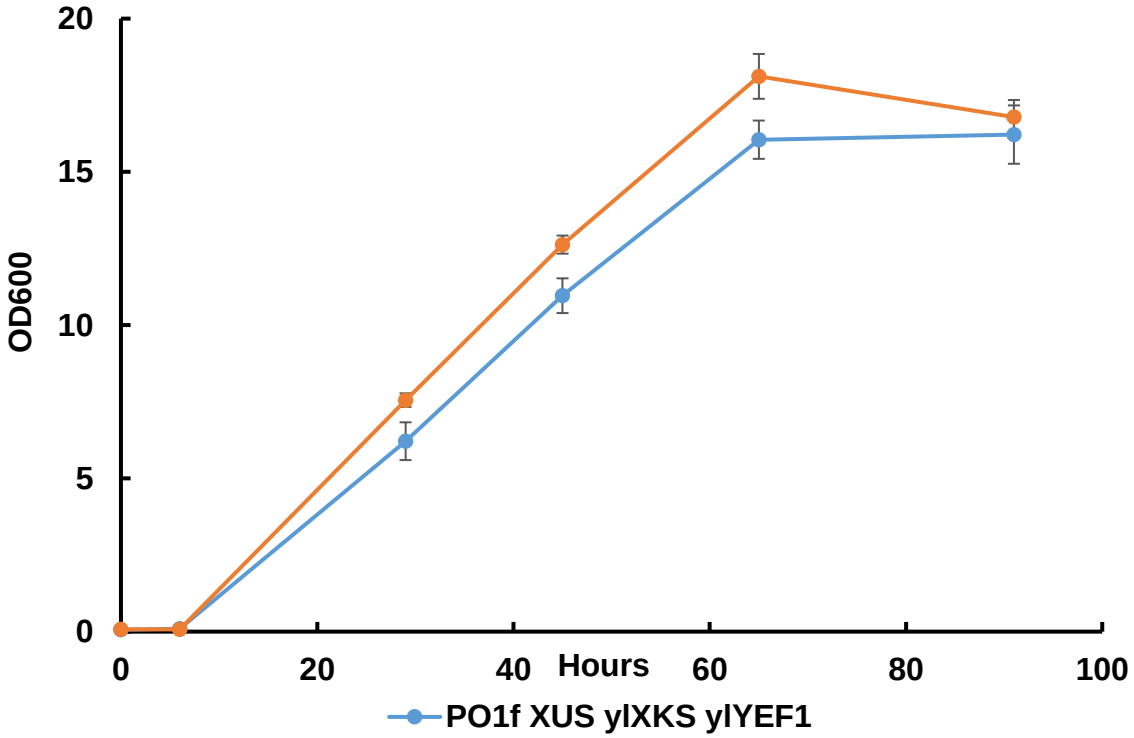


Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

3.2.2 Attempt to balance cofactors in xylose pathway by overexpression of ylYef1

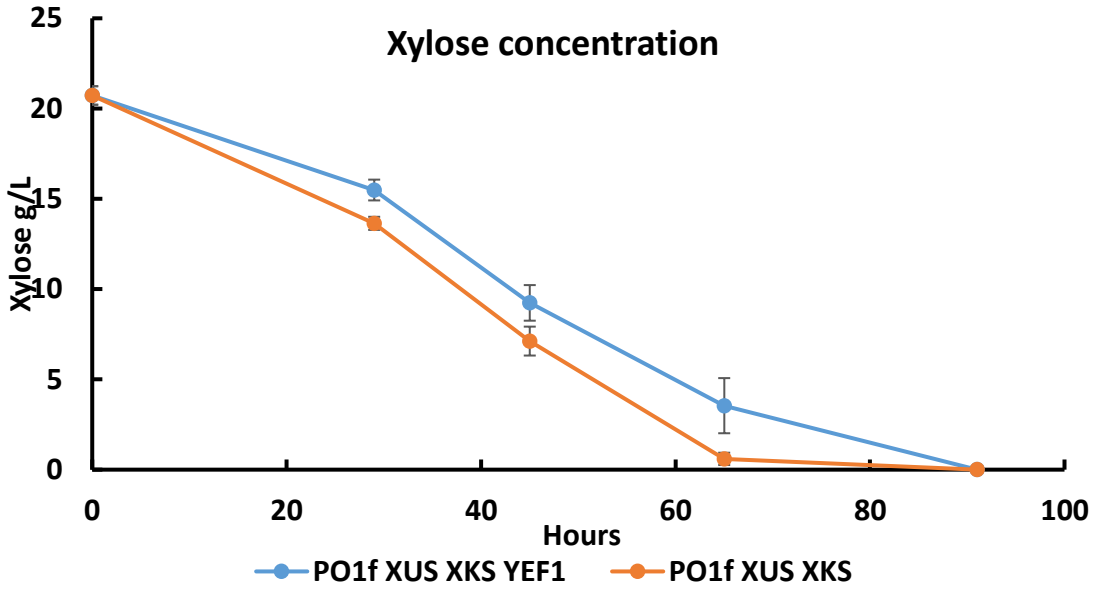
The xylose pathway includes a xylose reductase (XR) and xylitol dehydrogenase (XDH) in most fungi that can utilize xylose (Young, Lee and Alper 2010). The XR and XDH we previously introduced into *Y. lipolytica* was from *Scheffersomyces stipitis*. This XR, also known as XYL1 prefers NADPH as its co-factor and the XDR, known as XYL2 prefers NAD+ as a cofactor. There are three native NAD+/NADH kinases identified in *Y. lipolytica*. Among them, only ylYEF1 was identified to improve lipid yield (Qiao *et al.* 2015). However, none of these enzymes was measured for in vitro activity. In this study, we cloned ylYEF1 into an overexpression cassette and integrated it into PO1f XUS XKS strain. However, we were not able to observe an improvement in growth (Figure 3-4) or xylose utilization (Figure 3-5), and conversely, saw a slight drop in both. We also measure the lipid production for the final time point, and observed no visible difference between the two strains (data not shown).

Figure 3-4: Cell growth of PO1f XUS ylXKS strain overexpress ylYEF1 compared with PO1f XUS ylXKS as control.



Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

Figure 3-5: Xylose utilization of PO1f XUS ylXKS strain overexpress ylYEF1 compared with PO1f XUS ylXKS as control.



Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

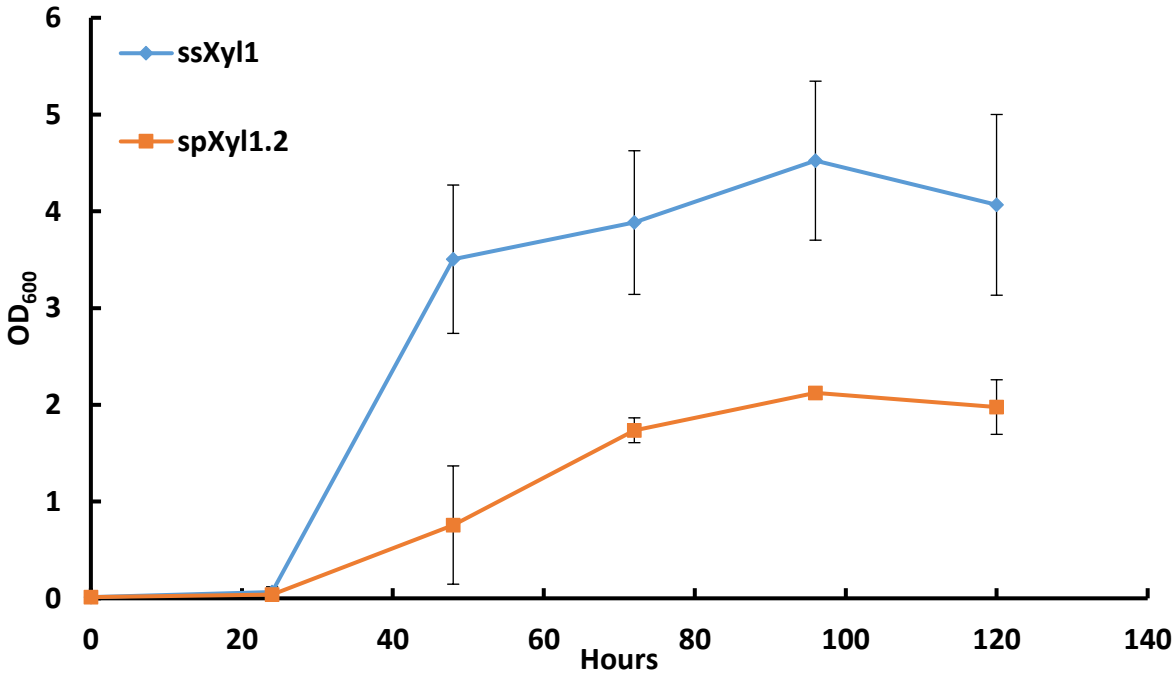
3.2.3 Attempt to balance cofactors in xylose pathway by overexpressing spXYL1.2

Most xylose reductase (XR) prefers NADPH over NADH (Bruinenberg *et al.* 1983). However, this preference can be reversed in some organism. *Spathaspora passalidarum* is an organism that grows very well on xylose and is the only organism without glucose inhibition on xylose utilization. Sequencing has demonstrated that this organism has two XRs. One of these enzymes prefers NADPH over NADH whereas the other prefers NADH over NADPH, which will be referred as spXYL1.2 in this chapter (Long *et al.* 2012; Cadete *et al.* 2016). The NADH preferred spXYL1.2 was expressed with strong promotor into the *Y. lipolytica* PO1f XUS XKS strain in an attempt to further improve xylose utilization in *Y. lipolytica.*

Since our previous xylose utilization strain has been starved to enable XYL1 and XYL2 to have multiple copies in the genome, this strain cannot be used to easily test spXYL1.2. Specifically, knocking out all copies of XYL1 in this strain would be very hard and reintegrating the new enzyme to have the same copy number of both XYL1 and spXYL1.2 is almost impossible. Thus, we rebuilt a strain with one copy of XYL2 and an extra copy of ylXKS overexpressed from strong hybrid promoters. Then we integrate one copy of ssXYL1 (the original enzyme used to generate the xylose utilizing strain) or spXYL1.2 into this strain to test their growths. This experimental design was chosen based on another study that showed overexpressing XYL1, XYL2 and XYL3 together can enable growth on xylose without starvation. We were able to observe growth with both versions of xylose reductase. However, the growth was much less than our starved

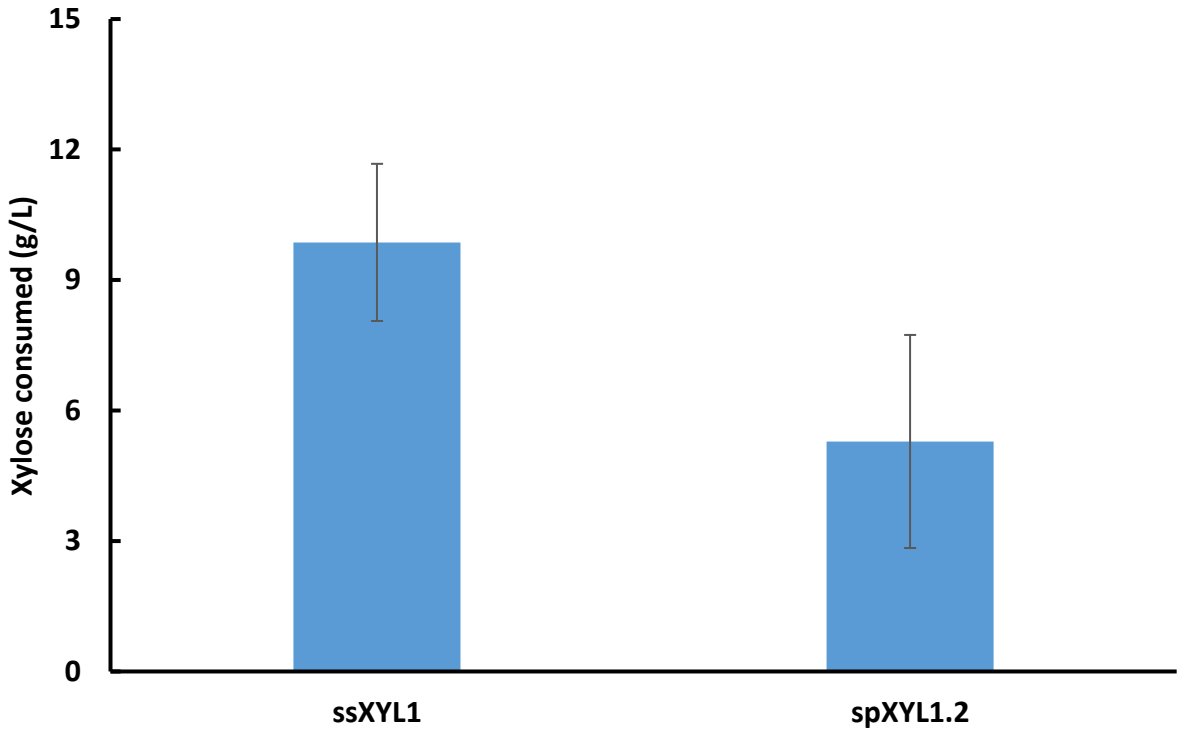
strain (Figure 3-6). Nevertheless, these strains still serve the purpose of comparing the two xylose reductases with different preference of cofactors (Figure 3-7).

Figure 3-6: Cell growth of PO1f Xyl2 ylXKS strain overexpress spXyl1.2 compared with PO1f Xyl1 Xyl2 ylXKS as control.



Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

Figure 3-7: Xylose usage of PO1f Xyl2 ylXKS strain overexpress spXyl1.2 compared with PO1f Xyl1 Xyl2 ylXKS as control.



Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

3.3 DISCUSSION

We tested seven targets in this chapter, and found only native xylulose kinase overexpression can improve xylose utilization. It is interesting that overexpression of the four targets in pentose phosphate pathway did not result in significant improvement, especially since these targets have been previously been tested in *S. cerevisiae* and shown to improve xylose utilization. Our hypothesis for this difference is that genes in the pentose phosphate pathway in *Y. lipolytica* are already been expressed at very high level (Wasylenko, Ahn and Stephanopoulos 2015b). Thus, none of these four enzymes are rate-limiting bottlenecks for xylose utilization. This hypothesis is supported by a previous study showing that *Y. lipolytica* diverse part of the carbon flux into pentose phosphate pathway to generate NADPH for lipid synthesis. Nevertheless, we determined that xylulose kinase is the bottleneck for xylose utilization and overexpressing native ylXKS can significantly improve xylose utilization. We also observed that overexpression of the native ylXKS not only improved cell growth on xylose, but also imparted a boost in cell growth on glucose. This similar growth improvement on glucose was not seen without the overexpression of XYL1 and XYL2. It is currently unclear what is causing this observed growth rate increase.

The several attempts to improve growth on xylose by balancing cofactors in *Y. lipolytica* was not successful. We think that the NAD+/NADH kinases ylYef1 may have depleted the NADH pool in *Y. lipolytica* and thus caused a drop in growth on xylose. Alternatively, the activity of spXyl1.2 may be too low in this host that the balancing of cofactors does not make up for reduced enzyme activity. As mentioned above, *Y. lipolytica* can generate NADPH by divert carbon flux into pentose phosphate pathway. In that case,

the xylitol accumulation we observed may not due to co-factor imbalance but from low activity of XYL2 compared to XYL1. To test this hypothesis, future study needs to introduce better xylitol dehydrogenase into *Y. lipolytica* to test whether xylitol accumulation can be reduced.

Chapter 4: Producing biochemicals in *Yarrowia lipolytica* from xylose through a strain mating approach

4.1 CHAPTER SUMMARY

Efficient xylose conversion is an essential feature for producing renewable biochemical from lignocellulosic hydrolysate. However, rewiring xylose catabolism is challenging, especially for unconventional yeasts. In this study, we demonstrate the efficacy of a yeast mating approach with *Y. lipolytica* that can combine a previously engineering and evolved xylose phenotype with a metabolite overproduction phenotype. Specifically, we mated several engineered *Y. lipolytica* strains that produce α-linolenic acid, riboflavin and triacetic acid lactone with an engineered and adapted xylose utilizing strain to obtain three diploid strains that rapidly produce these molecules directly from xylose. Specifically, titers of 0.52 g/L ALA, 96.6 mg/L riboflavin and 2.9 g/L TAL respectively from were obtained from xylose in flask cultures. This production level was generally on par or higher than the parental strain cultivated on glucose. This result showcases the first study using strain mating in *Y. lipolytica* for producing biomolecules from xylose, and thus demonstrates the utility of this approach as a routine tool for metabolic engineering.

4.2 INTRODUCTION

While xylose has long been considered an important carbon source for the production of renewable fuels and chemicals (Jeffries 1983; Kwak *et al.* 2019), enabling

full and efficient utilization remains a challenge in the field (Young, Lee and Alper 2010; Harner *et al.* 2015; Jin *et al.* 2015b; Sànchez Nogué and Karhumaa 2015). The underlying challenge is that xylose catabolism is non-native for most commonly explored microorganisms (Wang, Shopsis and Schneider 1980; Jeffries and Jin 2004; Jagtap and Rao 2018) and thus mismatched with production capacity. To this end, *Y. lipolytica* is a non-conventional, oleaginous yeast that can be rewired to produce a wide variety of products after metabolic engineering (Liu *et al.* 2015a, 2015c; Abdel-Mawgoud *et al.* 2018; Spagnuolo *et al.* 2018), but does not fastidiously utilize xylose (Wagner and Alper 2015). In the past few years, several labs have reported on engineered xylose utilization in *Y. lipolytica*, specifically for the production of lipids (Ryu, Hipp and Trinh 2015; Ledesma-Amaro *et al.* 2016; Li and Alper 2016; Rodriguez *et al.* 2016). However, conversion of xylose into alternative products has not been fully explored in this host partially due to the challenges implicit in genetically engineering *Y. lipolytica* (Larroude *et al.* 2018) and the reliance upon adaptive evolution as part of an engineering strategy in several cases. To address this limitation, we describe a yeast mating approach to combine xylose utilization capacity with biochemical overproduction into a diploid strain.

Strain mating is a traditional genetic approach that has been commonly used in both fundamental genetics studies and in industrial breeding programs (Steensels *et al.* 2014b). As examples, strain mating has been used breed *Saccharomyces cerevisiae* for a variety of desirable phenotypes including higher tolerance to ethanol, adaptation to cold temperatures, the ability to ferment oligosaccharides and the accumulation of more diverse flavor profiles (Sanchez, Solodovnikova and Wendland 2012; Steensels *et al.*

2014a, 2014b; Snoek *et al.* 2015; Gibson *et al.* 2017; Ogata, Iwashita and Kawada 2017). Haploid *Y. lipolytica* likewise have the ability to mate into diploid cells. However, unlike *S. cerevisiae* which genetically encode two sets of mating loci and possess the ability to mating type switch, *Y. lipolytica* only has a single mating locus and hence does not switch mating type. As a result, two mating types of *Y. lipolytica* exist, type A and type B, depending on what is coded in this singular locus (Barth and Gaillardin 1997). Within the *Y. lipolytica* mating locus is two genes: one that controls conjugation and the other that controls sporulation (Weber, Kurischko and Barth 1988; Barth 2011). While the genetics and physiology of mating have been explored in this host, very little application has been proposed and applied for strain breeding.

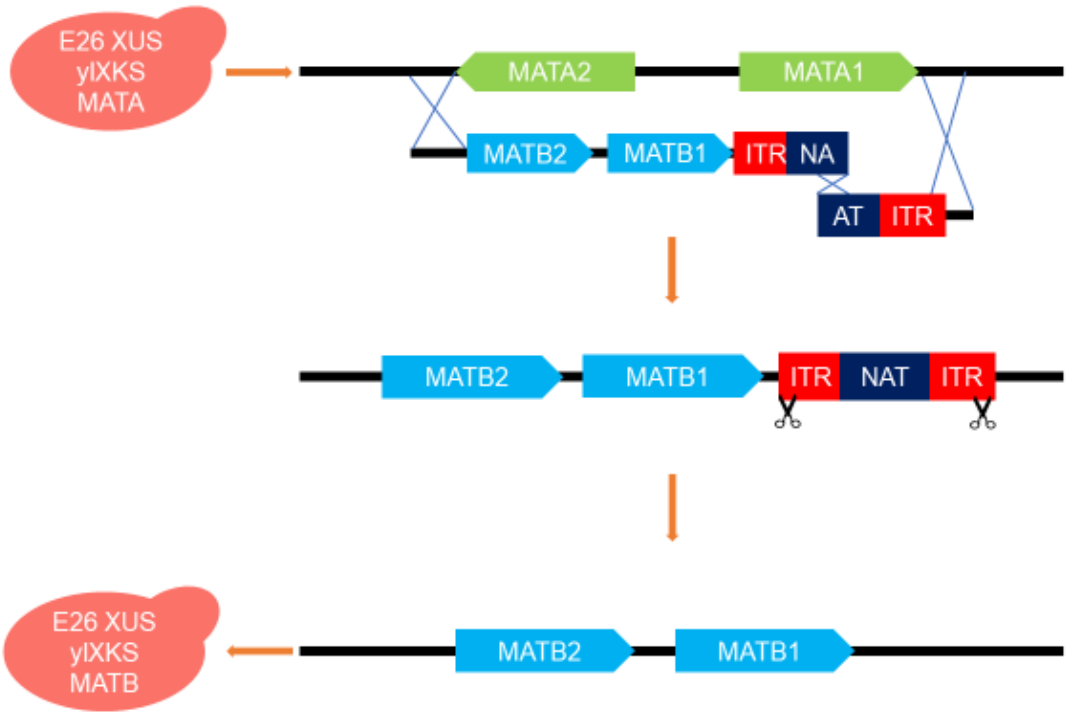
Here, we sought to utilize strain mating in *Y. lipolytica* to combine a xylose catabolic phenotype with biochemical overproduction. Specifically, we take a xylose utilizing strain (obtained partially via strain adaptation) and convert the mating type from type A to type B and then mated this strain with three other type A production strains (for α-linolenic acid, riboflavin, and triacetic acid lactone) to generate diploid strains capable of producing new metabolites from xylose. The results demonstrate the ability to retain both phenotypes from the haploid mated strains (i.e. high production and xylose utilization). In fact, production was generally higher from xylose. Collectively, these results show that strain mating can be a powerful tool to make diverse products from xylose in *Y. lipolytica*.

4.3 RESULTS

4.3.1 Converting the mating type of the xylose utilizing strain

The high lipid producing, xylose utilizing strain was engineered from the PO1f strain background (Beopoulos *et al.* 2008; Liu and Alper 2014; Li and Alper 2016), a mating type A strain. As this is a popular background host strain, we sought to mate-switch this singular strain to type B in an effort to serve as a viable mating partner to combine xylose utilization with any other type A strain. To do so, we utilized a piggyBac transposon system for a marker-less mating type switch (Wagner, Williams and Alper 2018). Specifically, *Y. lipolytica* wild-type MATB1 and MATB2 along with flanking DNA was synthesized by IDT and assembled into a singular plasmid together with the NAT dominant selection marker using Gibson assembly. The MATA1 and MATA2 regions was then replaced by the synthesized MATB1-MATB2-NAT construct via a split marker homologous recombination method. The NAT marker was flanked by a pair of piggyBac ITR sequences to enable excision by an excision-only piggyBac transposase mutant (Figure 4-1). Following PCR confirmation of the mating type switch, we used this resulting strain (derived from E26 XUS with the additional xylulokinase activity and mating type switch) for all subsequent mating studies. This mating type B strain will be referred as XUS-B as a short-hand for the rest of this paper.

Figure 4-1: Mating type switch of E26 XUS XKS strain.

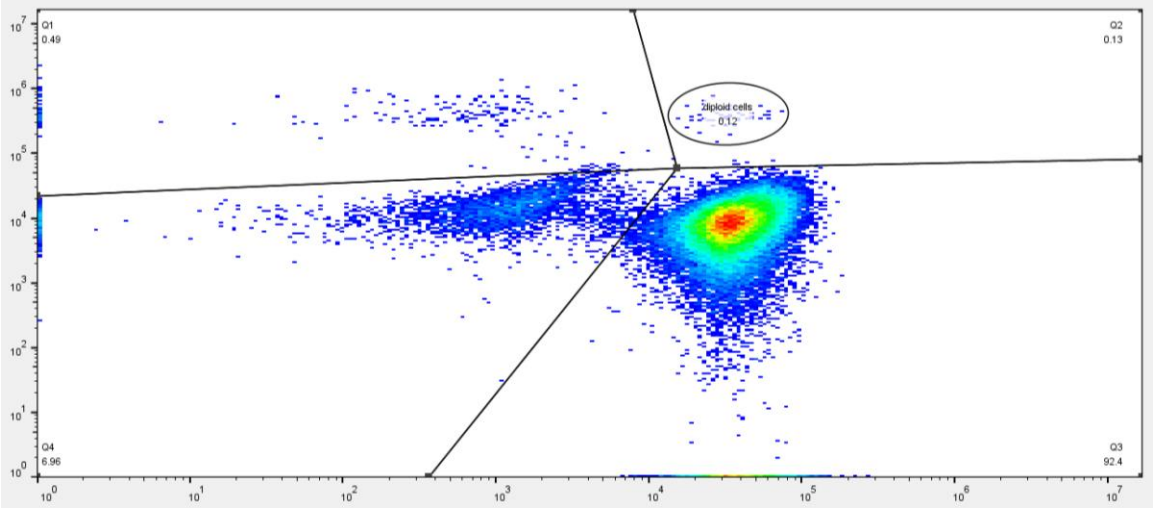


MATA1 and MATA2 region was first replaced by MATB1 and MATB2 and NAT marker with ITR flanking through homologous recombination. And then NAT marker was removed by excision-only piggyBac transposase.

4.3.2 Conjugation frequency

To quantitatively determine the conjugation frequency, we overexpressed far-red fluorescent protein mKate in XUS-B strain and mated it with a wild-type PO1f strain overexpressing humanized *Renilla reniformis*-derived GFP (hrGFP). The mated culture was quantified by a BD Accuri™ C6 Plus flow cytometer to determine the composition of haploid and diploid strains based on fluorescent protein expression (Figure 4-2). From this analysis, the conjugation frequency was estimated to be 0.081% within this unselected population, which is consistent with previous reports suggesting the frequency is 0.04% (Weber, Kurischko and Barth 1988). While this value is congruent with other literature values, we noted throughout this experiment that the absolute conjugation frequency can be highly variable and depends on the specific haploid strains used for mating.

Figure 4-2: Flow cytometry results of mated culture grow out in unselected media.

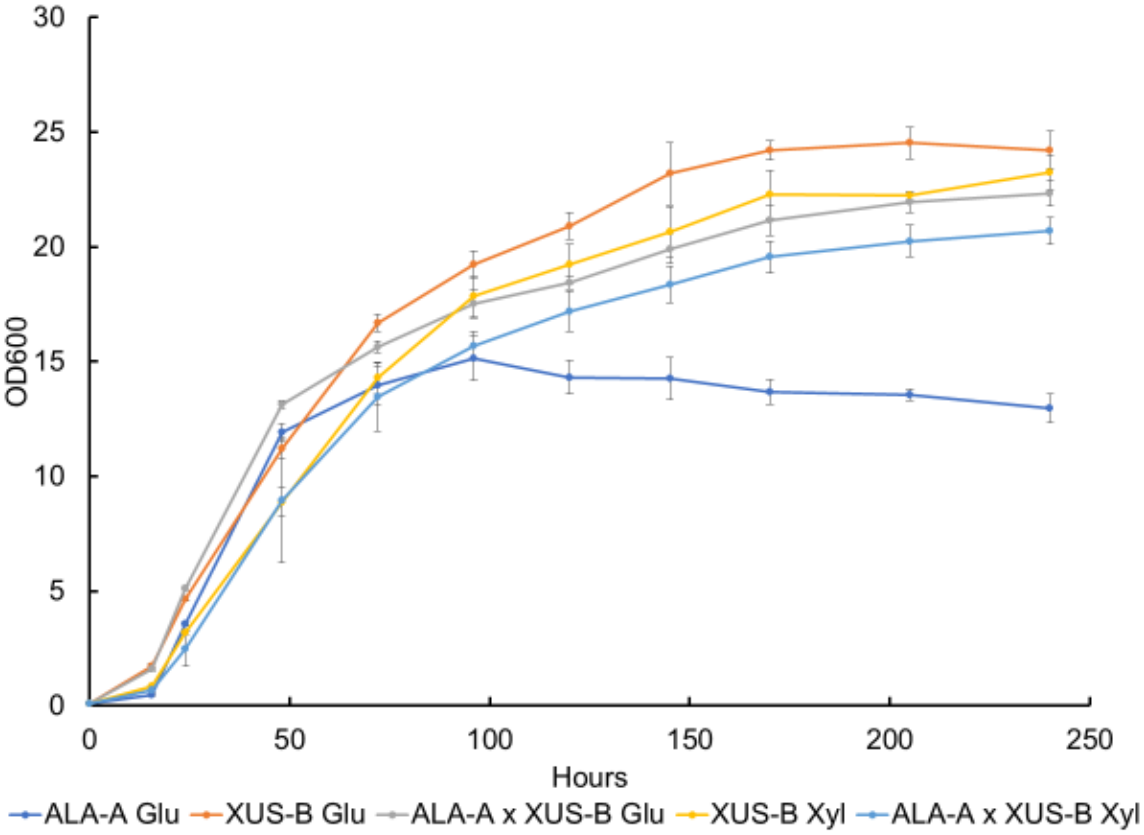


50,000 events were collected. Circled events represent diploid cells.

4.3.3 Achieving a xylose catabolizing strain overproducing ALA via mating

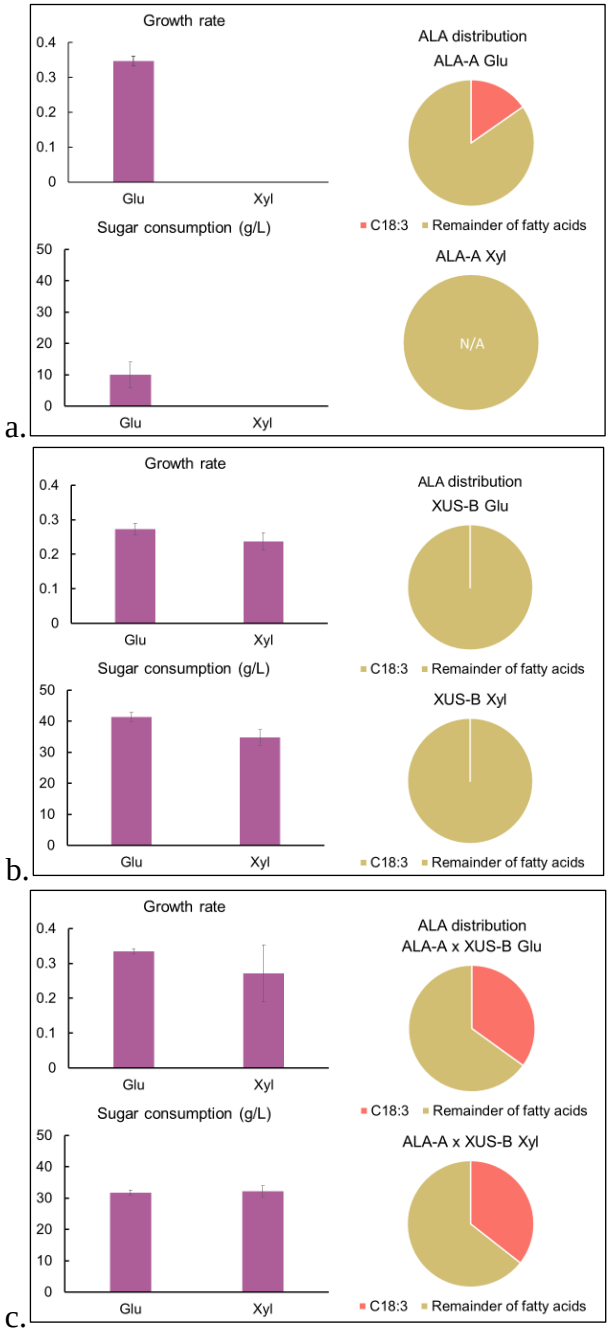
Given the oleaginous nature of XUS-B, we first decided to test the efficacy of creating a xylose-utilizing, α-linolenic acid producing strain (ALA-A). To do so, we chose our previously reported type A strain that produces α-linolenic acid (ALA) (Cordova and Alper 2018), a medically important polyunsaturated fatty acid with three double bounds (Narisawa *et al.* 1994; Stark, Crawford and Reifen 2008; Rodriguez-Leyva *et al.* 2010). Successful mating (following the protocol describe above with selection) resulted in diploid strains that produced ALA. The diploid strain XUS-B x ALA-A grows to a higher peak OD in both glucose and xylose (Figure 4-3) than the original ALA haploid strain. This mated, diploid strain also produces more total lipid when glucose is used as carbon source (1.74 g/L) and slightly less total lipid when xylose is the sole carbon source (1.47 g/L) than the original ALA strain produces in glucose (1.59 g/L). Within this lipid pool, the mated strain produced nearly 2.5 times more ALA (0.61 g/L) on glucose and more than 2 times more ALA (0.52 g/L) on xylose compared to the parental haploid strain, ALA-A (0.24 g/L). Percentage-wise, ALA comprised around 35% of total fatty acids from both glucose and xylose, whereas the haploid strain only had 15% of fatty acids as ALA on glucose in this run (Figure 4-4). These results demonstrate a more robust diploid strain with benefit realized in both glucose and xylose.

Figure 4-3: Cell growth of ALA, XUS and their mated diploid strain.



Growth curves of ALA-A, XUS-B and ALA-A x XUS-B strain flask experiment were shown with glucose or xylose as carbon source. Growth curve of ALA-A grow in xylose media is not shown due to lack of growth. Each data point represents an average ± 1 standard deviation of 4 biological replicate experiments that were grown in parallel.

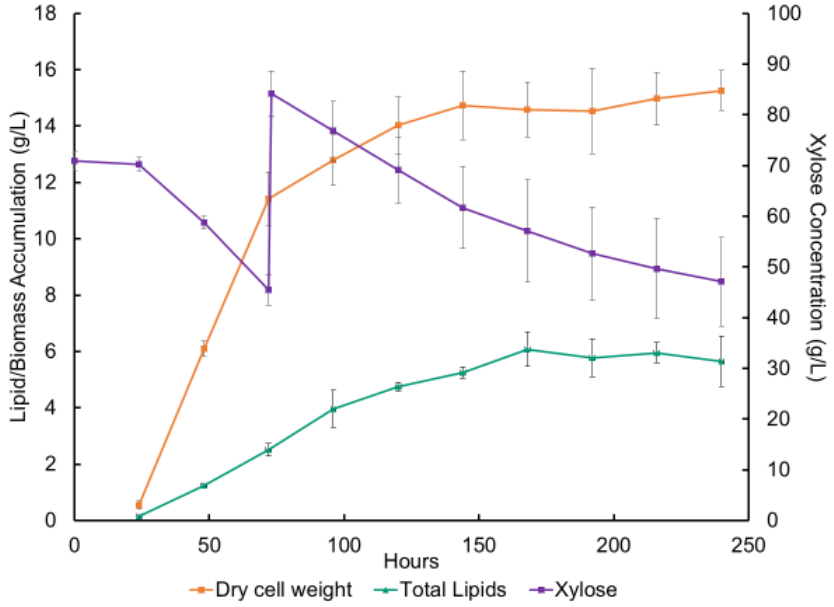
Figure 4-4: Growth rate, sugar consumption and ALA distribution of ALA-A. XUS-B and their mated diploid strain.



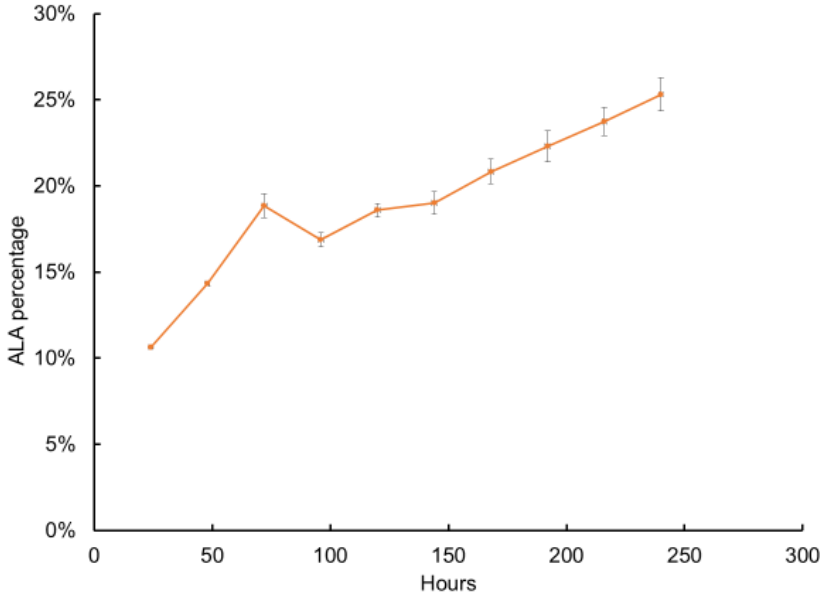
a. Maximum growth rate; b. sugar consumption and c. lipid distribution were measured at 240 hours of flask fermentation. Each data point represents an average ± 1 standard deviation of 4 biological replicate experiments that were grown in parallel.

Next, we evaluated ALA production of this diploid strain in 1.7 L bioreactors with xylose as the sole carbon source. To do so, we mimicked our previously-published bioreactor run and substituted xylose instead at the same concentration as glucose had been used. Similar to the performance of the parental strain on glucose, the diploid strain reached an optical density higher than 50 while grown on xylose. This reactor supported a dry cell weight of over 15 g/L, of which almost 6 g/L was lipid (Figure 4-5a). The final titer of ALA reached 1.42 g/L, equivalent to 25.3% of total lipid content (Figure 4-5b). While this number is lower than the 35.6% achieved in a flask setting, it is significantly higher than the 17% seen in the haploid strain grown on glucose under this same bioreactor setting. By the end of the run, a residual content of nearly 40 g/L of xylose remained and this strain exhibited a substantial decrease in oleic acid (C18:1) percentage (Figure 4-6).

Figure 4-5: Bioreactor fermentation of ALA-A x XUS-B strain.



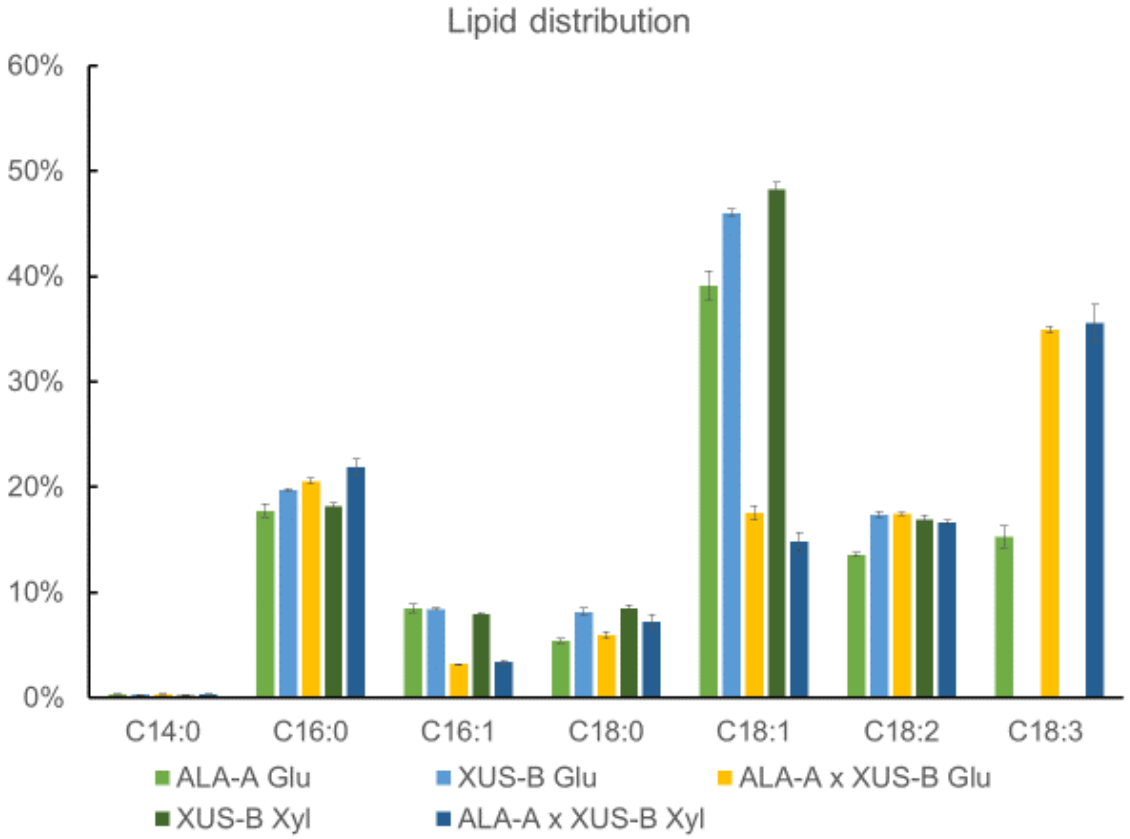
a.



b.

a. Dry cell weight, total lipid and xylose concentration were determined from duplicate analysis of biological replicates. b. ALA content and ALA percentage of total lipid were determined from duplicate analysis of biological replicates. Each data point represents an average ± 1 standard deviation of 2 biological replicate experiments that were grown in parallel.

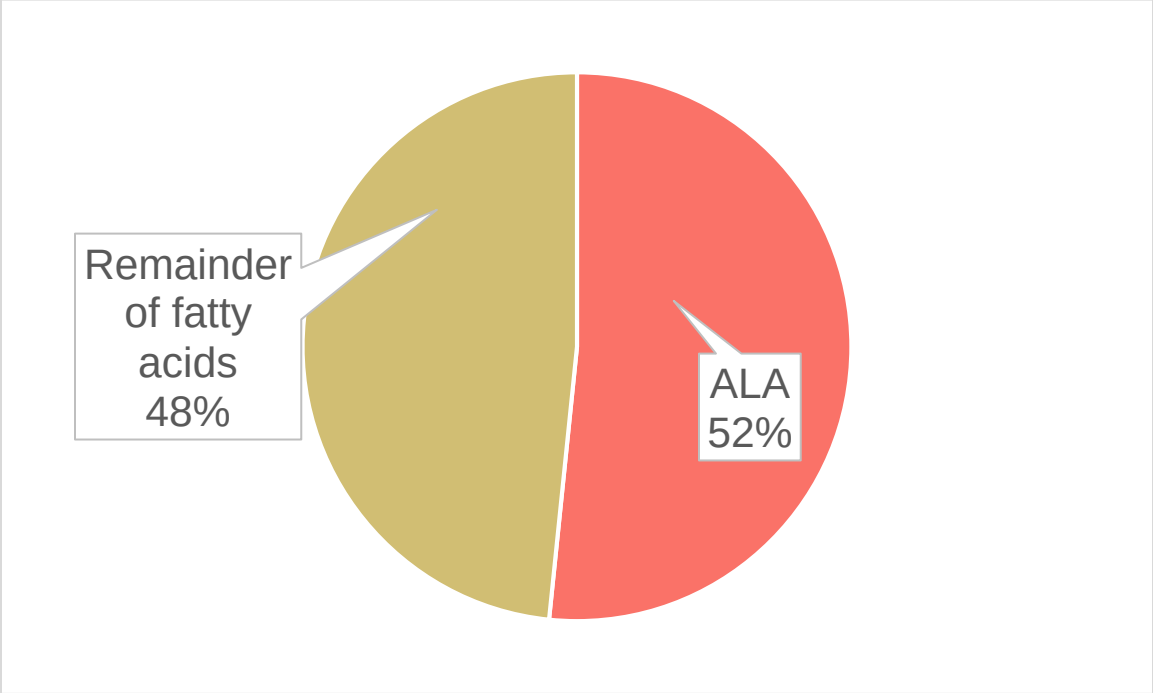
Figure 4-6: Lipid distribution at 240 hours in flask experiment.



Lipid distribution of ALA in xylose media is not shown due to lack of growth. Each data point represents an average ± 1 standard deviation of 4 biological replicate experiments that were grown in parallel.

The performance of this diploid ALA strain obtained by crossing ALA-A x XUS-B achieved high content of ALA approaching, if not ever so slightly exceeding, the 33% threshold suggestive of multiple positioning on the triglyceride backbone. To follow up on this result, we measured ALA production in xylose-containing rich medium (YP+xylose) and determined that ALA actually exceeded 50%, indicating that xylose utilization (in complex media) can enable ALA to be substituted on multiple backbone positions (Figure 4-7). As a comparison, the most common source of dietary supplemented ALA is from flaxseed oil, and the ALA percentage in flaxseed oil is around 58% (Bozan and Temelli 2008). The ALA content in our diploid Y. lipolytica is very close to flaxseed oil. Furthermore, the ALA percentage within the bioreactor increases after total lipid has stabilized, thus indicating the potential that the desaturase enzyme can function not only on free fatty acid but also on fatty acid chains on TAG.

Figure 4-7: Lipid distribution of ALA-A x XUS-B strain grown in YPX media.



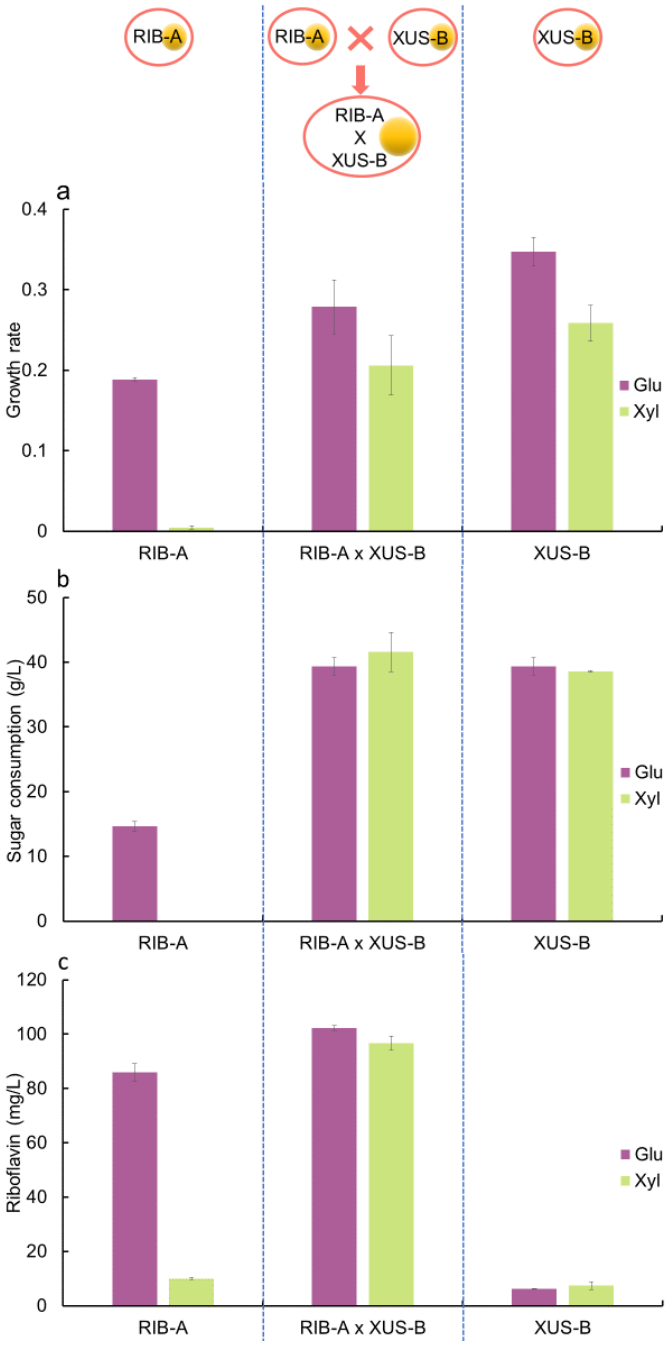
The result represents an average of 4 biological replicate experiments that were grown in parallel.

4.3.4 Mating XUS-B with a riboflavin producing strain

Next, we sought to mate the XUS-B strain with a previously reported riboflavin (vitamin B2) producing strain (Wagner *et al.* 2018). Riboflavin is a dietary vitamin commonly produced by microorganisms industrially (Schwechheimer *et al.* 2016; Revuelta *et al.* 2017). Mating the riboflavin producing strain with a xylose utilizing strain has the benefit of increasing flux through the pentose phosphate pathway so that it could optimally feed into riboflavin production (Schwechheimer *et al.* 2016).

To create a xylose-catabolizing, riboflavin producing strain, we mated the XUS-B strain with a metabolic engineered riboflavin production strain termed RIB-A (Wagner *et al.* 2018). This haploid, parental RIB-A strain had already undergone multiple rounds of mutagenesis and selection to enable more than 85 mg/L extracellular riboflavin in defined medium with 40 g/L initial glucose at 400 hours (Figure 4-8c). The mated diploid strain of RIB-A x XUS-B grow on xylose and produced 102 mg/L riboflavin from glucose, and 97 mg/L from xylose, both of which exceed the value from the parental RIB-A strain grown on glucose (Figure 4-8 a-c). In contrast, the other parental strain (XUS-B) only produces 6.2 and 7.4 mg/L riboflavin using glucose and xylose as carbon sources, respectively.

Figure 4-8: Growth rate, sugar consumption and riboflavin production of RIB-A, XUS-B and their mated strain in flask experiment.

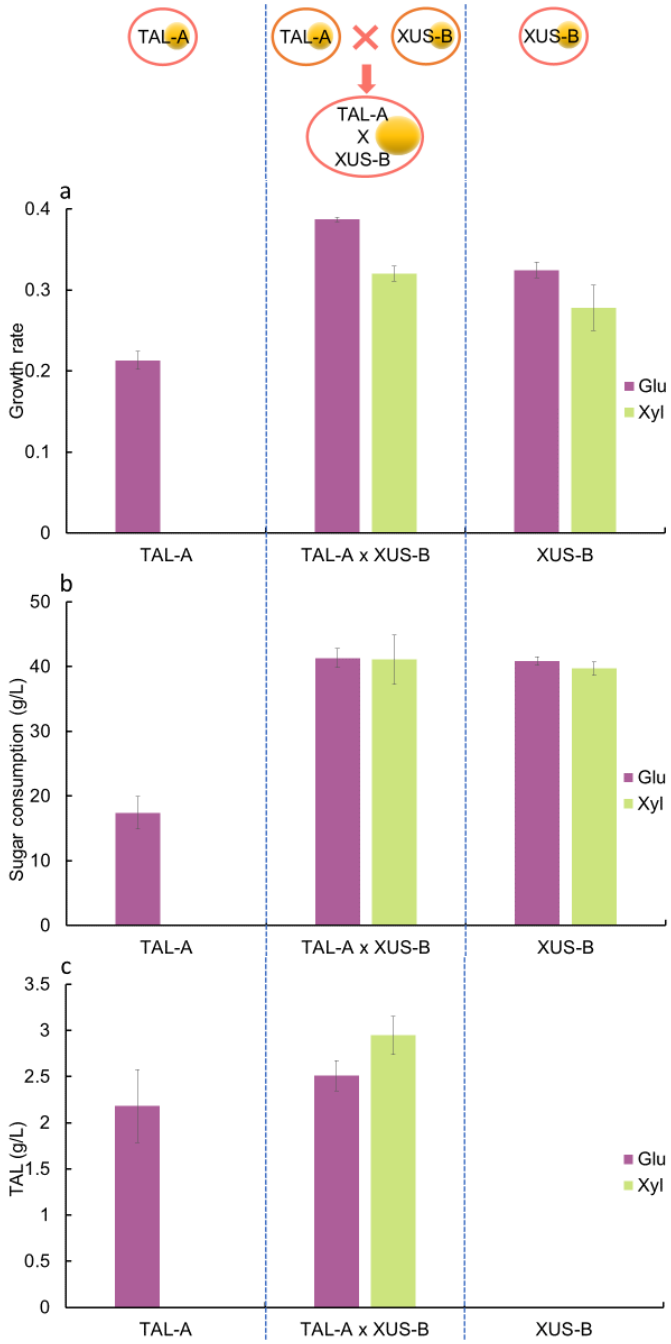


a. Maximum growth rate; b. sugar consumption at 400 hours; c. riboflavin concentration at 400 hours. Each data point represents an average ± 1 standard deviation of 3 biological replicate experiments that were grown in parallel.

4.3.5 Mating with TAL strain

As a final demonstration, we chose to create a xylose-utilizing polyketide strain. Polyketides are an interesting class of secondary metabolites with great potential (Robinson 1991; Lim, Go and Yew 2016; Palmer and Alper 2019). For this study, we chose to mate our XUS-B strain with a rewired strain of *Y. lipolytica* engineered to overproduce substantial amounts of triacetic acid lactone (TAL) (Markham *et al.* 2018). Here, the mating type A TAL production strain (TAL-A) when mated with XUS-B was able to achieve high growth, reaching OD levels of 24.9 and 20.9 on glucose and xylose compared with a value of 10.2 for the parental TAL producing strain (TAL-A). In similar fashion to the riboflavin strain, this diploid strain more rapidly and completely consumed sugars during the flask fermentation (Figure 4-9b) when compared with the parental TAL-A strain. Whereas the TAL production was faster initially from the TAL-A strain, final fermentation analysis shows that the diploid strain produced 2.6 and 2.9 g/L TAL from glucose and xylose, respectively compared with the 2.3 g/L produced from the haploid strain (Figure 4-9c). Finally, we measured lipid content in the cells and saw that only 0.23 g/L of lipid was produced in the diploid strain, thus this imparted phenotype did not strongly adversely impact TAL production.

Figure 4-9: Growth rate, sugar consumption and TAL production of TAL-A, XUS-B and their mated strain in flask experiment.



a. Maximum growth rate; b. sugar consumption at 168 hours; c. TAL concentration at 168 hours. Each data point represents an average ± 1 standard deviation of 4 biological replicate experiments conducted from separate mating events that were grown in parallel.

4.4 DISCUSSION

This study demonstrates the feasibility of rapidly and efficiently enabling xylose catabolism in previously-engineered *Y. lipolytica* strains. Remarkably, the diploid strains could sustain production from xylose that were on par or higher than levels from glucose. In the end, we demonstrated production of 0.52 g/L ALA, 96.7 mg/L riboflavin and 2.9 g/L from xylose respectively. The parity of glucose- and xylose-based production confirms the strong carbon flux through the PPP in this organism.

Overall, we observed that diploids tended to be more robust with respect to overall growth than haploids—a phenomenon seen in *S. cerevisiae* as well (Galitski *et al.* 1999). Specifically, we found that overexpression of the native ylXKS not only improved cell growth on xylose, but also imparted a boost in cell growth on glucose. This similar growth improvement on glucose was not seen without the overexpression of XYL1 and XYL2. While it is currently unclear what is causing this observed growth rate increase, it can be seen in haploid and diploid strains with this genotype. In fact, the poorer growth seen in all three of the haploid production strains was revived by the mating (Figure 4-4, 4-8a-b, 4-9a-b). Specifically, all diploid strains more thoroughly used sugars in these fermentations (with a substantial amount going to biomass). We posit that the increased growth rates are also advantageous in the stability of the strains. In particular, diploid yeast can return to haploid form through sporulation. For *Y. lipolytica*, sporulation has been reported to be triggered by citrate sporulation medium or commercial V8 juice (Gaillardin, Charoy and Heslot 1973; Barth and Weber 1985), although sporulation has also been observed in the YM mating media itself. In previous reports, sporulated culture

will have a brownish color, which we did not observe. Likewise, attempts to sporulate these strains were not successful using techniques reported in the literature. Prior reports also suggest that *Y. lipolytica* can go through parasexual cycle to convert from diploid to haploid (Kurisciiko 1986). To evaluate this phenomena, we isolated colonies after fermentations here and did not find any haploid cells emerging in the culture. These findings suggest that the strong stability (and robust growth) of these mated strains are attractive for industrial processes. To further test the robustness of the mating approach, the TAL strain mating event and resulting flask fermentation was conducted using replicates separate mating events. The resulting production was very consistent across these separately mated diploid strains, another positive marker for strain robustness/stability.

Each of the diploid strains isolated in this work took advantage of xylose utilization and antibiotic resistance already in the type A strain. It is possible that alternative uses of this technology may not have an easily selectable genotype. In such cases, due to the low efficiency of mating, it would be necessary to establish a rapid method for identifying metabolite production. It is likely that further improvements to the mating process itself could further increase efficiency. Finally, the type B strain used in this work was a high lipid producer. Such a phenotype is not always compatible with the end goal (as was the case with TAL production as these lipids divert flux away from the TAL pathway). As a result, future extensions of this work could create a type B strain that is not a lipid-hyperproducer.

This study demonstrates the efficacy of *Y. lipolytica* mating for strain engineering efforts and expanding the product portfolio from xylose. This approach allows for a far quicker approach to strain engineering than importing the xylose catabolic pathway and undergoing strain adaptation. Thus, mating with pre-optimized strains can be beneficial in this host.

4.5 CONCLUSION

In summary, this study shows that *Y. lipolytica* mating can be used as a powerful tool for strain engineering in this host. The mated strains reached titers of 0.52 g/L ALA, 96.6 mg/L riboflavin or 2.9 g/L TAL respectively from xylose in flask conditions. These titers and resulting growth and sugar consumption rates indicates that diploid strains as generated here are quite fit and robust. The use of a mating approach in *Y. lipolytica* can rapidly combine two previously engineered phenotypes.

Chapter 5: Enabling glucose/xylose co-transport in yeast through the directed evolution of a sugar transporter[2][2]

5.1 CHAPTER SUMMARY

The capacity to co-transport glucose and xylose into yeast has remained a technical challenge in the field. While significant efforts have been made in transporter engineering to increase xylose transport rates, glucose-based inhibition still limit most of these transporters. To address this issue, we further engineer sugar transporter proteins to remove glucose inhibition and enable glucose/xylose co-transport. Specifically, we start with our previously derived CiGXS1 FIM mutant strain and subjugate it to several rounds of mutagenesis and selection in a hexose metabolism null strain. Through this effort, we identify several mutations including N326H, a truncation in the C-terminal tail, I171F, and M40V as additionally dominant for reducing glucose inhibition. The resulting transporter shows substantially improved xylose transport rates in the presence of high quantities of glucose including up to 70 g/L glucose. Moreover, the resulting transporter enables co-utilization of glucose and xylose with glucose rates on par with a wild-type transporter and xylose rates exceeding that of glucose. These results demonstrate that major facilitator superfamily hexose transporters can be rewired into

[2] Content in this chapter adapted from a previously authored publication written by HL. HL equally contributed to the contributed to experiments, analyses and writing of the manuscript.
Haibo Li, Olivia Schmitz and Hal S. Alper, 2016. Enabling glucose/xylose co-transport in yeast through the directed evolution of a sugar transporter. *Applied Microbiology and Biotechnology* 100(23):10215-10223.

glucose-xylose co-transporters without functional inhibition by either substrate. These results enhance the potential of using lignocellulosic biomass as a feedstock for yeast.

5.2 INTRODUCTION

First generation biofuels and biochemicals derived from food crops have significant economic and societal drawbacks (Naik *et al.* 2010). As a result, the use of second generation biofuels produced from lignocellulosic biomass has been studied extensively (Lynd 1996; Naik *et al.* 2010; Dale and Ong 2012). Depending on the organism used, several challenges exist with respect to the monosaccharide profile resulting from hydrolysate (Kumar et al. 2009; Chiaramonti et al. 2012; Brethauer and Studer 2015; Marriott et al. 2015). Of high importance for the common bioethanol producing yeast *Saccharomyces cerevisiae* (Van Vleet and Jeffries 2009), lignocellulose hydrolysate is comprised of not just glucose, but also xylose which can vary from 5% to 35% of total sugar content depending on the biomass source (Gírio *et al.* 2010). Since wild-type *S. cerevisiae* strains are unable to natively catabolize xylose (Jeffries and Jin 2004; Sànchez Nogué and Karhumaa 2015), significant metabolic and evolutionary engineering strategies have been applied to improve utilization (Jeffries 1983; Kuyper et al. 2003; Lee et al. 2012, 2014; Scalcinati et al. 2012). As a result, xylose catabolic rates have been improved in recent years using a variety of strategies (Jeffries and Jin 2004; Salusjärvi *et al.* 2006; Young, Lee and Alper 2010; Scalcinati *et al.* 2012; Shen *et al.* 2012; Kim *et al.* 2013; Parreiras *et al.* 2014; Harner *et al.* 2015; Bamba, Hasunuma and Kondo 2016).

Despite success in improving xylose catabolic rates, *S. cerevisiae* still suffers from efficient co-utilization of glucose and xylose due to a strong native preference for glucose as the primary carbon source (Matsushika *et al.* 2009; Young, Lee and Alper 2010). In a typical fermentation, xylose is utilized only when glucose is completely consumed, which leads to additional decreases in efficiency due to higher concentrations of ethanol and acetic acid that inhibit xylose catabolic rates (Klinke, Thomsen and Ahring 2004; Demeke *et al.* 2013). As a result, a renewed interest has been placed on enabling co-utilization of these primary carbon sources in lignocellulosic biomass.

A number of previous studies suggest that a primary factor driving glucose preference originates at the sugar transporter proteins in the cell membrane (Salusjärvi et al. 2006, 2008; Subtil and Boles 2012). Specifically, while some hexose transporters are able to facilitate xylose transport, they are typically strongly inhibited by glucose even at low concentrations (Sedlak and Ho 2004; Young *et al.* 2011). Several groups have studied the potential to engineer such transporter proteins for improved rates of xylose transport (Farwick *et al.* 2014b; Nijland *et al.* 2014; Shin *et al.* 2015; Wang, Yu and Zhao 2015). Previously, our lab identified a sequence motif in symporters that can rewire glucose transporters (such as the *Candida intermedia* glucose-xylose symporter, CiGXS1 (Leandro, Gonçalves and Spencer-Martins 2006)) into a xylose exclusive transporter (Young et al. 2012, 2014). However, these transporters were still typically inhibited by glucose (esp. at relevant concentrations in liquid culture) and thus were not fully sufficient for co-utilization applications.

In this study, we further engineer sugar transporter proteins to remove glucose inhibition and enable glucose/xylose co-transport. Specifically, we start with our previously derived CiGXS1 FIM mutant strain and subjugate it to several rounds of mutagenesis and selection. This selection took place in a mutant strain of yeast devoid of all glucose transporter proteins and also containing deletion in key hexokinase enzymes. As a result, it was possible to use glucose as an "inhibitor" during the directed evolution process. After several rounds of evolution, we identify a glucose/xylose co-transporter that can tolerate high glucose concentration and can co-transport both sugars at equal concentrations with a preference for xylose. In doing so, this engineered transporter further enables the capacity for co-fermentation of lignocellulose derived sugars.

5.3 RESULTS

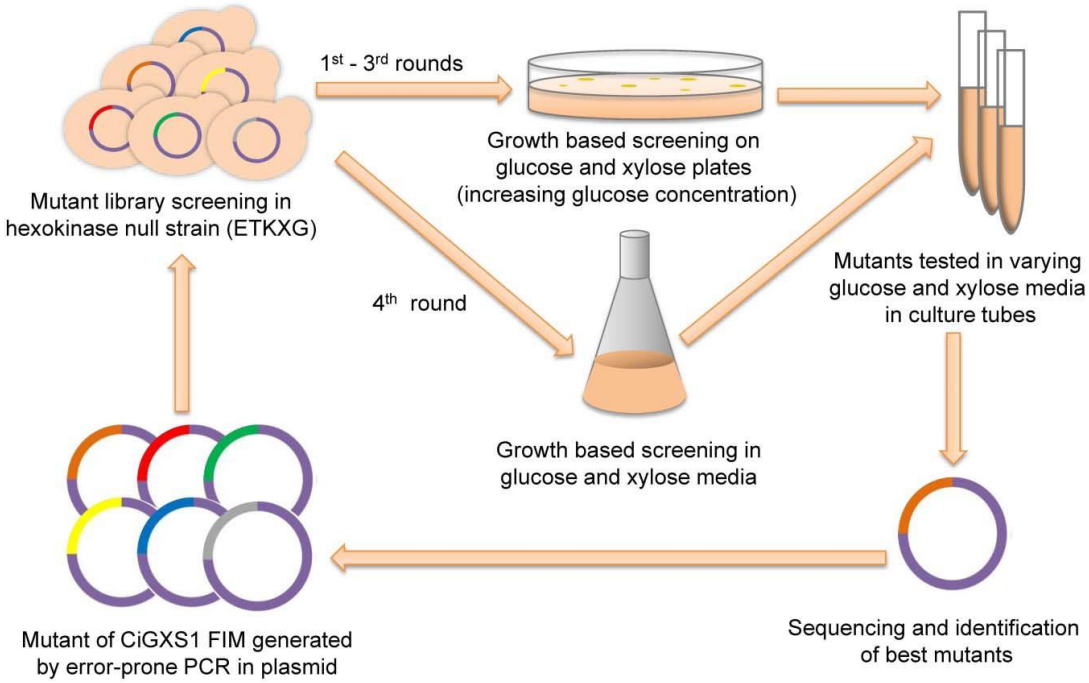
5.3.1 Constructing a platform screening strain for reduced inhibition

In previously work, we identified that a specific sequence (namely V38F L39I F40M) within the variable region of the G-G/F-XXX-G motif of transporters like CiGXS1 can attenuate glucose transport without compromising xylose transport capacity (Young *et al.* 2014). However, xylose transport was still highly inhibited by glucose for this CiGXS1 FIM transporter and similarly designed transporters. Therefore, we sought to reduce the glucose inhibition here through several rounds of directed evolution. To do so, it was necessary to create a screening strain that could enable selection of xylose

transporters that were uninhibited by glucose. Specifically, we started with the EBY.VW4000 strain of *S. cerevisiae* that has 20 deletions to remove all hexose (and hence pentose) sugar transport (Wieczorke *et al.* 1999). This host was used in our previous work to select for xylose transporter function and evolution (Young *et al.* 2014).

To establish a growth-based screen for transporters uninhibited by glucose, it was necessary to inactivate glucose metabolism in this host strain. To do so, we deleted key hexokinase steps in this background to create a *hxk1*, *hxk2*, and *glk1* null strain, in similar fashion to a recent report (Farwick *et al.* 2014b). Next, xylose utilization genes *XYL1* and *XYL2* from *Scheffersomyces stipitis* as well as *GAL2* were cloned onto plasmids and transformed to create the strain ETKXG. This resulting strain was shown to be able to grow on xylose and galactose, but not glucose or maltose. As a result, this platform strain is capable of being utilized in growth-based selections for transporters that enable xylose transport in the presence of increasing concentrations of glucose. The degree of inhibition can be selected by controlling the concentration of glucose used during the evolution. Thus, this strain was used for mutant selection throughout this study. The design and selection process is illustrated in Figure 5-1.

Figure 5-1: Design strategy of library construction and selection.

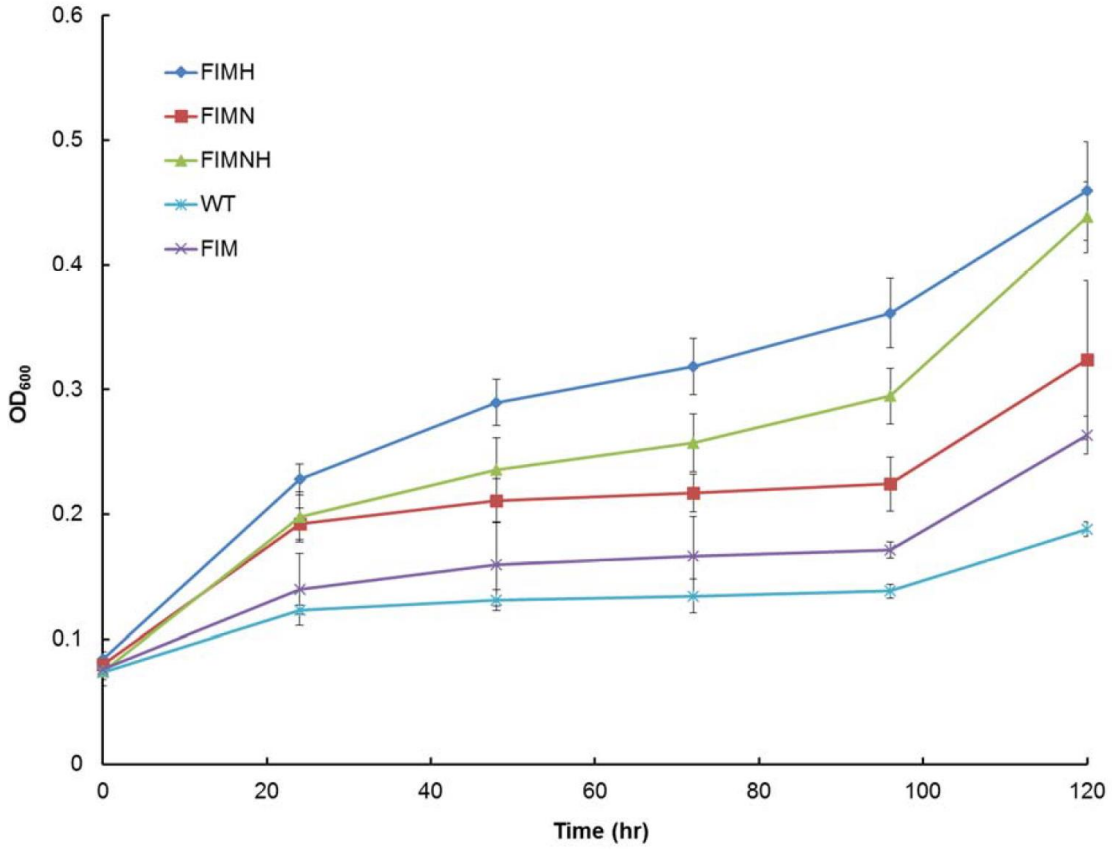


Mutant transporters were constructed by error-prone PCR and tested in hexokinase null strain (ETKXG). Libraries derived in rounds 1 through 3 were selected on plates with increasing glucose concentration. Libraries in the 4th round of selection were subcultured in flasks. Initially identified mutants were further tested in test tube fermentation and the best mutants were sequenced and used as the template for the next round of mutagenesis.

5.3.2 N326H and T170N were identified to reduce glucose inhibition in a first round of selection

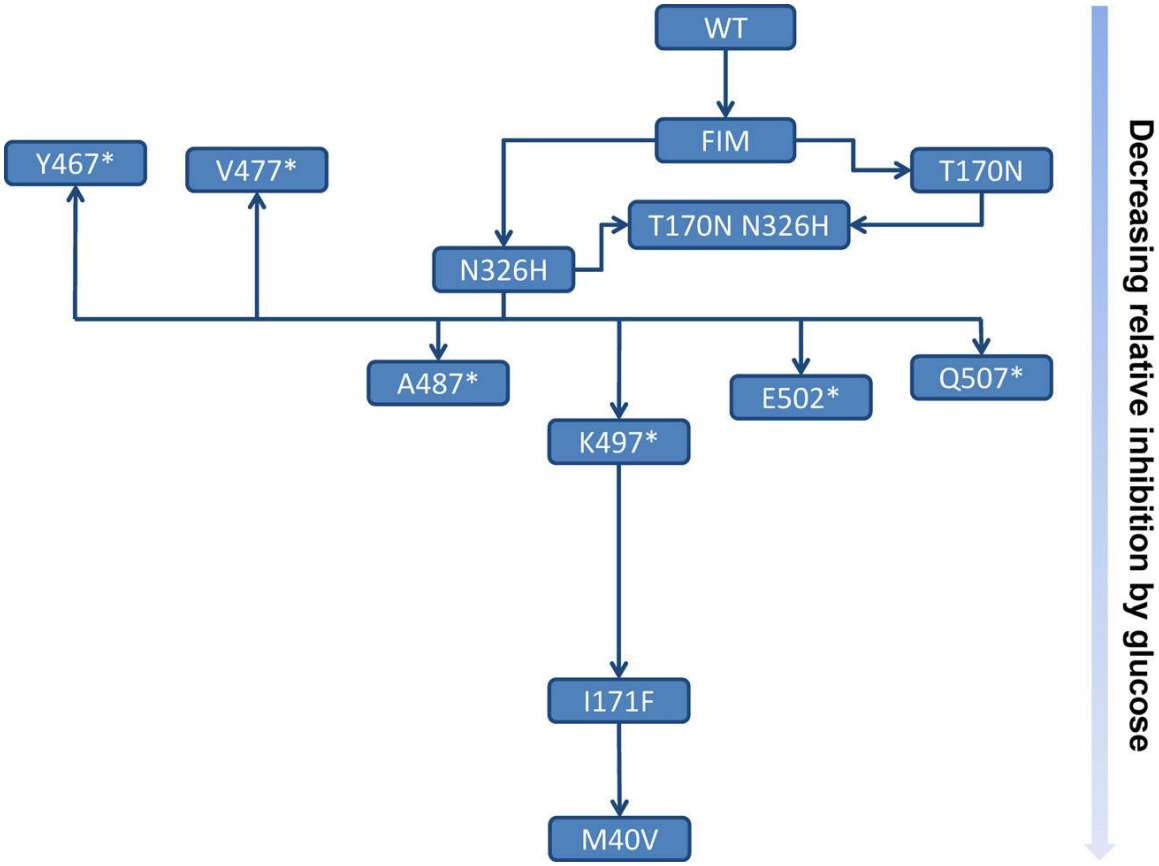
To reduce the glucose inhibition present in the CiGXS1 FIM transporter, we created error prone-PCR based libraries of this transporter protein followed by selection in the strain described above. In this first round of selection, yeast cells were selected on plates containing 20 g/L glucose and 20 g/L xylose. Such a concentration imposed a high selection pressure on the library and resulted in few recovered colonies. A total of eight mutants were isolated during this first round that showed reduced inhibition to glucose than the starting transporter, CiGXS1 FIM. Among these eight mutant strains, half had mutations to residue N326 and the remaining half had mutations to T170. For the transporters showing mutations at N326, three were changed to histidine and the other to serine. We observed that the histidine mutation had a more reduced inhibition than the serine mutation. All of the mutations to residue 170 changed the threonine to an asparagine. A corresponding mutation has been identified in *gal2*, *hxt5*, *hxt7*, *hxt11* and *hxt36* (Farwick *et al.* 2014b; Nijland *et al.* 2014; Shin *et al.* 2015; Reider Apel *et al.* 2016). CiGXS1 FIM N326H mutant (CiGXS1 FIMH) was less inhibited by glucose than T170N (CiGXS1 FIMN) (Figure 5-2). The entire strain selection process is shown depicted in Figure 5-3, starting from strain CiGXS1 FIM with level of glucose inhibition plotted across the y-axis. Table 5-1 lists all of the positive mutations found during each round of selection.

Figure 5-2: Growth curve of ETKXG expressing CiGXS1 FIMH, FIMN and FIMNH compared to CiGXS1 WT and FIM.



Each data point represents an average value ±1 standard deviation calculated from the results of 6 biological replicates.

Figure 5-3: Phylogeny of mutant transporter strains identified in this study.



Key mutations identified across the evolutionary trajectory explored here are shown. Arrows between boxes indicate the evolutionary trajectories taken in this work. The placement of boxes aligns with the 48-h glucose inhibition level in 40 g/L glucose and 20 g/L xylose media, with mutations on the top demonstrating more inhibition by glucose than those on the bottom.

Table 5-1. Positive mutations identified during the all four rounds of mutagenesis. A table cataloguing the mutations identified in this study is provided.

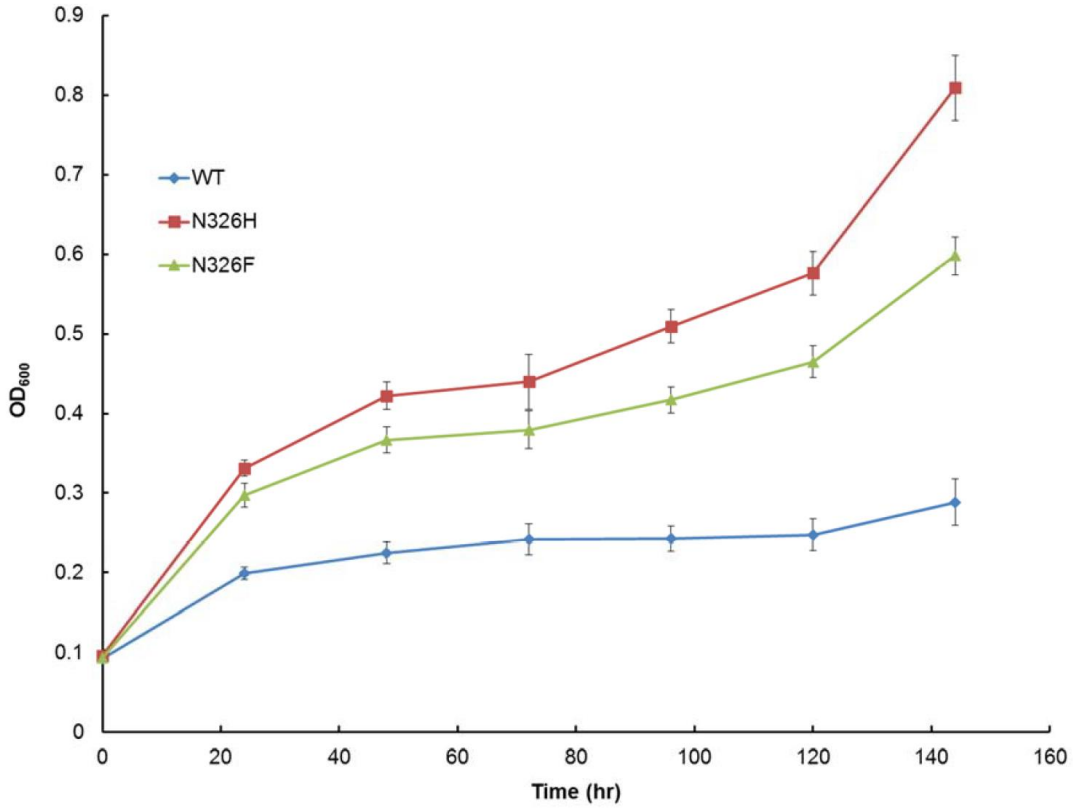
	Parental Strain	Mutations	Silent mutations
First round	FIM	N326H	
		N326H	
		L196F, N326H, M345T	V29V, L247L
		D70V, T170N, E231V, K275E	R193R, H277H
		T170N	A197A, N455N
		T170N	
		N326S, V504M	S177S, L201L
		T170N	T284T
Second round	FIMH	T361S	
		T361A, F500V	
		C89W, T361A, A366S, A431V, D439G	
		V377I	
		I27F, Q507*	
		F174L	
		F388V	
		A462T	
		S329F	L213L
		T170N, H326N, S368T, E502*	
		F67Y	L57L
		V138I, K181G	S163S
		L175F	
		A462T, D509G	
		L175F, N370I	
		F174L	
		T354S	I205I
		K511E	
		Q507*	

Table 5-1 continued. Positive mutations identified during the all four rounds of mutagenesis. A table cataloguing the mutations identified in this study is provided.

Third round	FIMH497	I171F	
		I205V, T284I	S72S, P240P
		I362F	S72S, P240P
		L108I	
		K440Q	Y55Y, S222S
		E17V, V132I, L407M	C128C, V271V, G328G, L340L, L346L, L423L
		I171N, T187S	
		G341D	
		V400I, T409S	T170T, A387A
		F385Y, T428I	
		V377I	N112N
		H185N, E258D	
		K228E	L79L, L236L
		A20P, L407F	
		V377I	N112N
		M353I, G452S	I157I, P406P
		S276P	D242D
		Q226H	
Fourth round	FIMH497	M40V, A160V	

Since two dominant mutations were identified in this first round, we sought to combine these mutations to test for additive phenotype effects. Specifically, we constructed a double mutant transporter CiGXS1 FIMNH with both T170N and N326H mutations. The double mutant FIMNH is less inhibited by glucose than FIMN alone, however, it is more inhibited than just FIMH alone. Finally, the N326H mutation corresponds to a similar residue modification found in a recent report (Farwick *et al.* 2014b). Specifically, this recent report demonstrates that changing asparagine 376/370 to phenylalanine for *gal2* and *hxt7* resulted in xylose transporters with decreased glucose inhibition (Farwick *et al.* 2014b). The mutation found in this study (N326H) was compared with the N326F suggested by the previous study. By doing so, we demonstrate that for CiGXS1, N326H is less inhibited by glucose than N326F (Figure 5-4). With this novel mutation validated and tested against prior reports, we sought to further reduce glucose inhibition by using CiGXS1 FIMH as a template for a second round of mutagenesis and selection.

Figure 5-4: Growth curve of ETKXG expressing CiGXS1 N326H and N326F compared to WT CiGXS1.



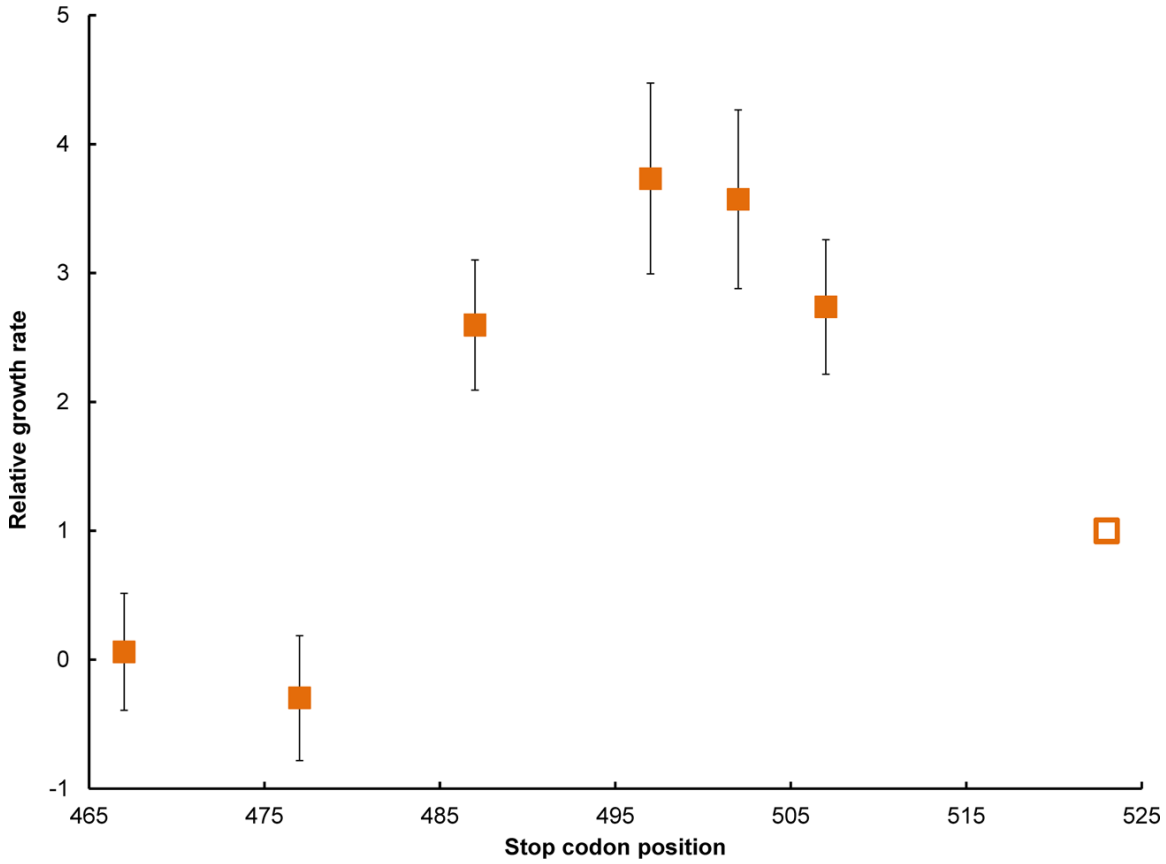
Each data point represents an average value ±1 standard deviation calculated from the results of 6 biological replicates.

5.3.3 A truncation in the transporter C-terminus further reduces glucose inhibition

Following a similar strategy as outlined above, a new mutant library was generated using CiGXS1 FIMH as a template and was selected on enhanced glucose conditions of 40 g/L glucose and 20 g/L xylose plates. After this selection, a total of 19 mutants were isolated that exhibited reduced glucose inhibition over FIMH (Figure 5-3, Table 5-1). Once again, a dominant mutation pattern arose. Specifically, six of the mutants were found to have mutations in the C-terminal tail of the transporter with three of these resulting in a truncated protein due to a stop codon. This result strongly suggested that a truncation of the C-terminal tail could decrease glucose inhibition.

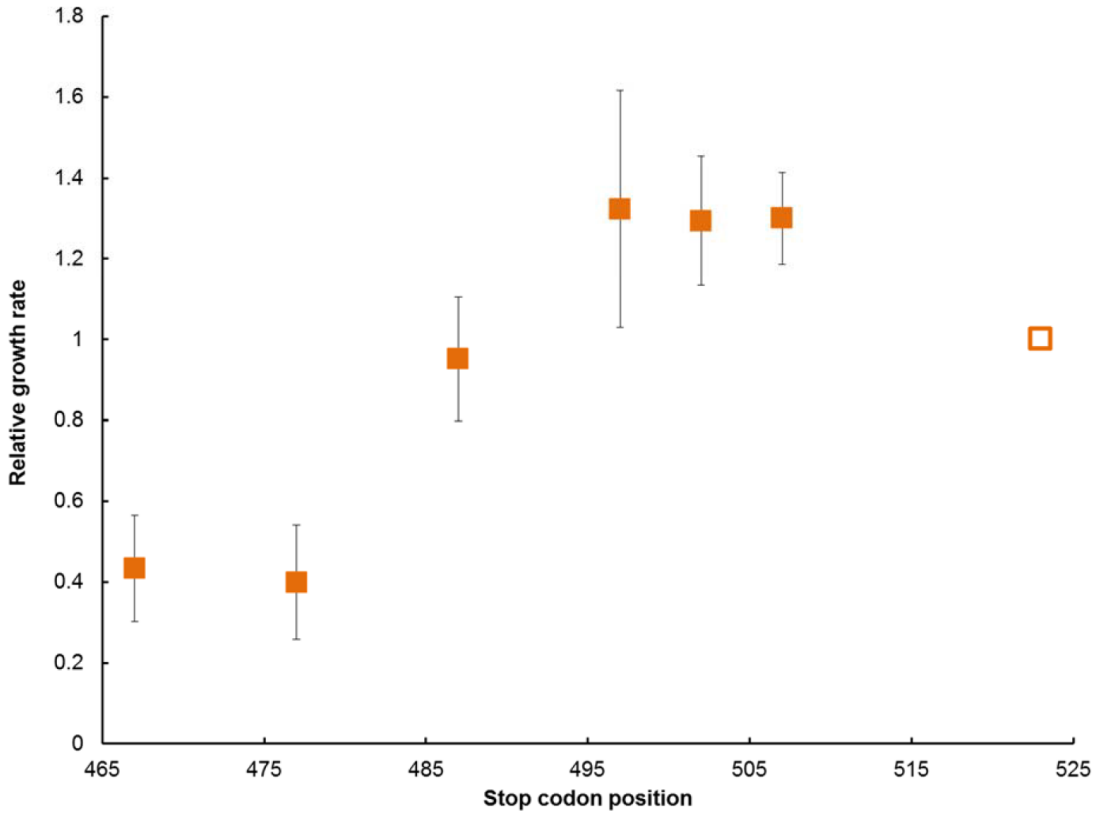
To further test this hypothesis and more systematically find an optimal truncation locus, we created a truncation library consisting of seven different truncations on the C-terminal tail of FIMH. Specifically, residue 467, 477, 487, 497, 502 and 507 were replaced with stop codons. Figure 5-5 shows the growth rate for CiGXS1 FIMH truncation mutants growing in 50 mL flasks with 40 g/L glucose and 20 g/L xylose relative to the growth of the full-length, 523 amino acid transporter. These results illustrated that a truncation at residue 497 yielded a transporter with the least degree of inhibition in the presence of glucose. This mutant was termed CiGXS1 FIMH497* and served as a template for the next round of mutagenesis. To test the generic nature of these findings, we created a similar truncation series in the CiGXS1 FIMN background and once again found position 497 as being optimal for reduced inhibition (Figure 5-6).

Figure 5-5: Relative growth rate of ETKXG strain expressing CiGXS1 FIMH with different C-terminal tail truncations.



Filled squares represent mutants with early termination. The empty square represents the original, full-length transporter. Each data point represents an average value ±1 standard deviation calculated from the results of four biological replicates.

Figure 5-6: Relative growth rate for ETKXG expressing CiGXS1 FIMN with different C-terminal truncations.



Filled squares represent mutants with early termination. Empty square represents original C-terminus tail length. Each data point represents an average value ±1 standard deviation calculated from the results of 4 biological replicates.

5.3.4 Two additional mutations (I171F and M40V reversion) further reduce glucose inhibition

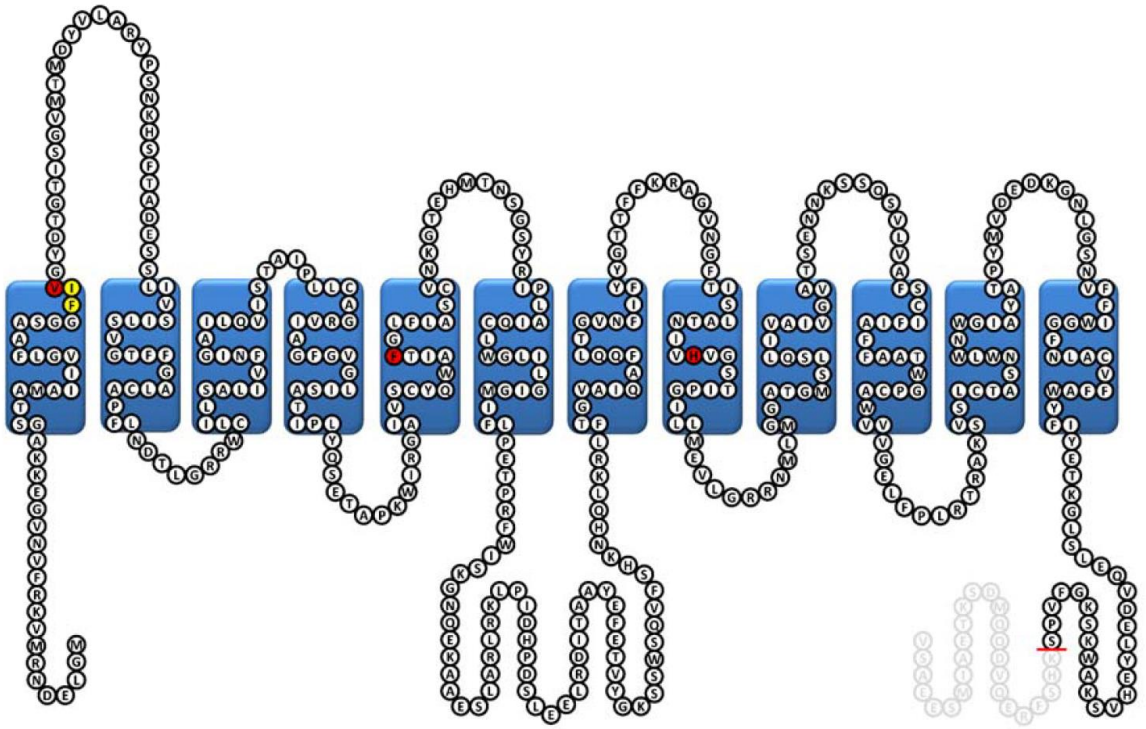
Finally, we conduct two additional rounds of mutagenesis and selection to further reduce glucose inhibition. First, the CiGXS1 FIMH497* truncation mutant identified above was used as the template for the third round of mutation and selection. This library was selected on 60 g/L glucose and 20 g/L xylose and resulted in 18 mutations demonstrated reduced inhibition (Table 5-1). While no multiple occurrences of mutations were observed in this round, the most improved phenotype resulted from the mutation I171F. It is interesting to note that the I171 residue is adjacent to the T170 residue identified in the first round of mutation. Moreover, while N326H and T170N did not have an additive effect (Figure 5-3), the N326H and I171F combination is better than N326H alone.

The resulting mutant, CiGXS1 FIMFH497* was then used as a template for a final round of mutagenesis and enrichment on 60 g/L glucose and 20 g/L xylose in liquid culture with three serial subcultures. In the end, we observed a singular mutation in the selected population. Interesting, the mutation was on M40 where methionine was converted to valine. This mutation belongs to the G-G/F-XXX-G motif that led to the starter transporter for this study (Young *et al.* 2014). This reversion of methionine to valine allowed for a restoration of glucose transport, but retained the lack of inhibition from glucose.

5.3.5 Collective comparison of xylose transporters demonstrates a reduction of glucose inhibition

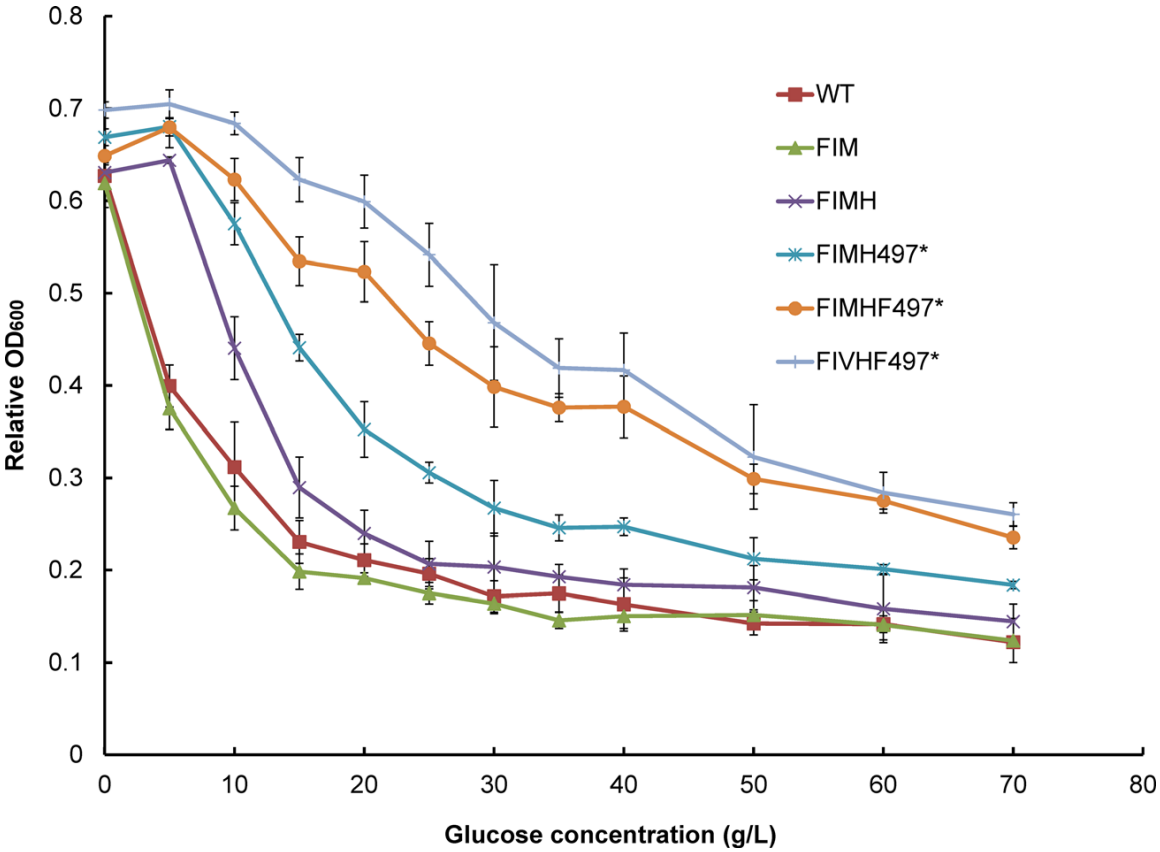
Figure 5-7 illustrates the overall mutations found across all rounds of mutagenesis and selection leading to the FIVFH497* transporter mutant. Each of the main mutations across this evolutionary trajectory were cultured in a deep well fermentation experiment simultaneously with 20 g/L xylose and varying amounts of glucose ranging from 0 g/L to 70 g/L. Figure 5-8 highlights the cell growth after 168 hours as a function of glucose concentration. This data clearly illustrates that as each round of evolution progressed, significant improvements were achieved in reducing overall glucose inhibition. Figure 5-9 shows a similar trend with respect to the absolute growth rate of the culture. This glucose inhibition level (as calculated by growth rate in the first 48 hours in 40 g/L xylose and 20 g/L glucose media) was used as the basis for ordering and spacing the mutations on Figure 5-3. Moreover, it is important to note that in addition to the reduced glucose inhibition, these transporters exhibited improved xylose transport rates (as highlighted by the growth in 0 g/L of glucose). This data demonstrates that CiGXS1 FIVFH497* is a strong xylose transporter with significantly attenuated glucose inhibition and retained xylose rates.

Figure 5-7: Predicted sequence–structure of CiGXS1 FIVFH497*.



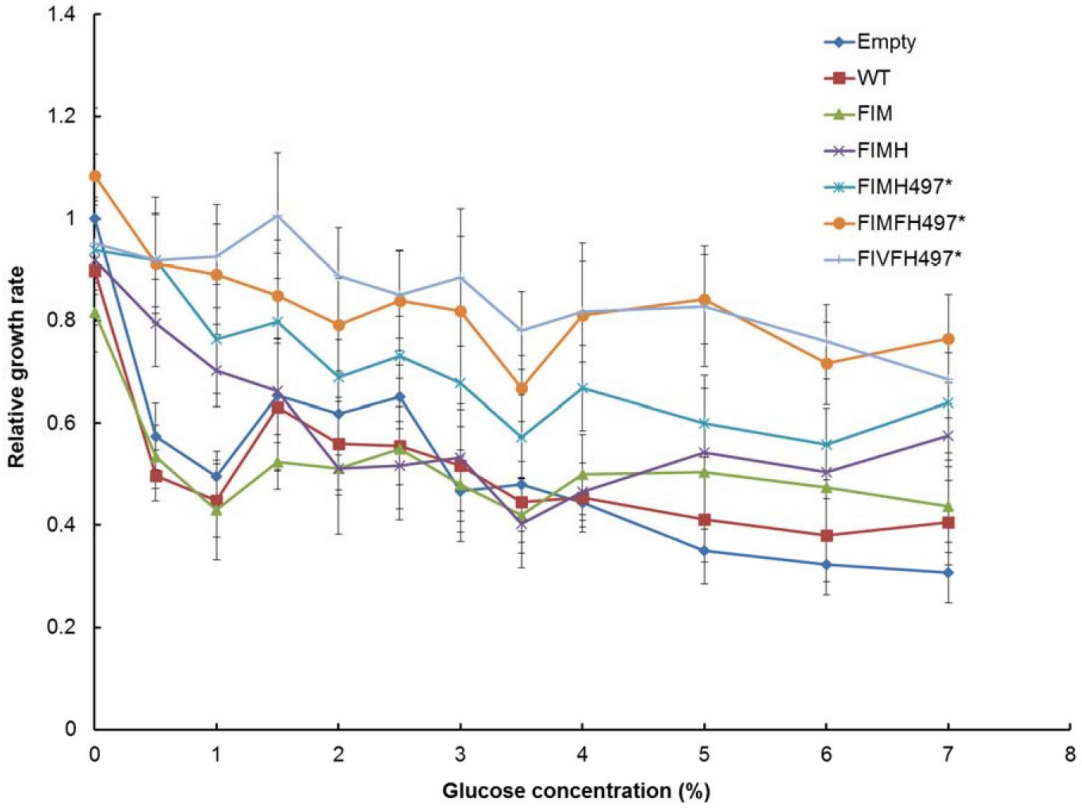
Mutations identified in this study are colored in red, mutations established by our previous studies are colored in yellow. The optimized C-terminal truncation site is shown with red line. The truncated C-terminus tail region is shown in grey. Transmembrane helix domains, helix tail domains and loop structures are predicted by HMMTOP with the extracellular region shown on top and cytosol shown below.

Figure 5-8: Relative growth levels of the ETKXG strain expressing various CiGXS1 mutants.



Deep well plate fermentations were conducted with all strains. OD600 was taken at 168 h of fermentation. For this study, 0, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, and 70 g/L glucose was used along with 20 g/L xylose in all cases. Each data point represents an average value ±1 standard deviation calculated from the results of four biological replicates

Figure 5-9: Relative growth rate of ETKXG expressing CiGXS1 mutants.

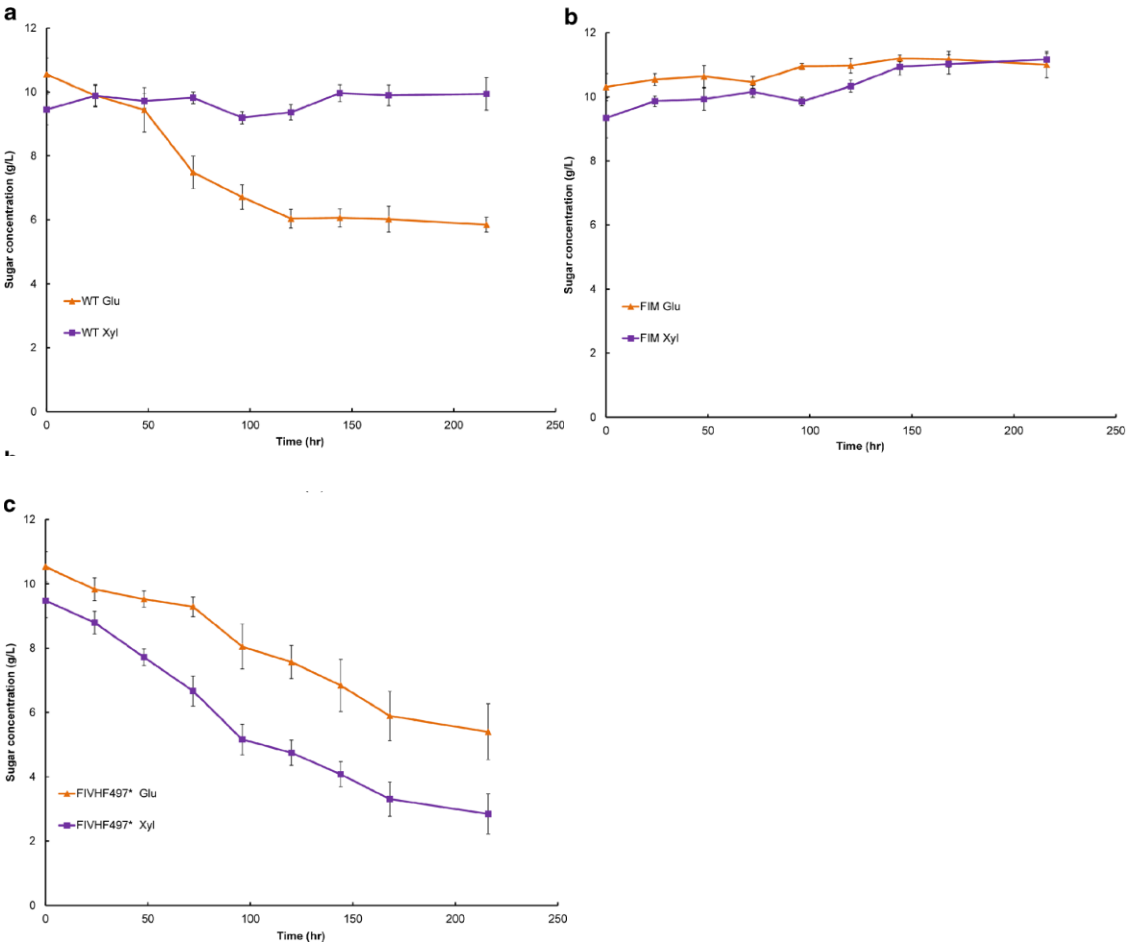


Deep well plate fermentations were conducted with all strains. OD600 was taken at 168 hours of fermentation. For this study, 0 g/L, 5 g/L, 10 g/L, 15 g/L, 20 g/L, 25 g/L, 30 g/L, 35 g/L, 40 g/L, 50 g/L, 60 g/L, and 70 g/L glucose was used along with 20 g/L xylose in all cases. Each data point represents an average value ±1 standard deviation calculated from the results of 4 biological replicates.

5.3.6 CiGXS1 FIVFH497* enables yeast to co-transport glucose and xylose

Owing to the final mutation in this selection process, it was possible that the CiGXS1 FIVFH497* could be a potent transporter that enables co-utilization of sugars. To test for such co-utilization, the parental EWV4000 strain outfitted with the *XYL1* and *XYL2* genes from *S. stipitis* were transformed with CiGXS1 WT, FIM and FIVFH497* and grown in a culture containing 10 g/L glucose and 10 g/L xylose. Sugar concentrations were monitored throughout the flask fermentation. Figure 5-10a demonstrates that the wild-type transporter was able to transport glucose to a point and xylose utilization was completely inhibited. Figure 5-10b highlights our prior results in which the CiGXS1 FIM is no longer able to transport glucose, but xylose transport is inhibited by the presence of glucose. Finally, the CiGXS1 FIVFH497* was tested and Figure 5-10c shows that this transporter mutant enables co-utilization of both sugars. The glucose transport rate observed here is on par with the rate observed in the wild-type transporter and the xylose transport rate exceeds that of the glucose transporter rate. This observed ability to transport xylose at high efficiency while maintaining the ability to transport glucose makes this transporter a superior candidate for co-fermentation applications. Collectively, these results demonstrate the complete rewiring of a predominantly glucose symporter into xylose/glucose co-transporter.

Figure 5-10: Glucose and xylose cofermentation.



EBY.VW4000 expressing CiGXS1 WT (4a), FIM (4b), and FIVFH497* (4c) were grown in 10 g/L glucose and 10 g/L xylose media in shake flask. Both glucose and xylose concentrations were measured at various times in the fermentation. Each data point represents an average value ±1 standard deviation calculated from the results of six biological replicates

5.4 DISCUSSION

Co-utilization of glucose and xylose (at relevant substrate concentrations in liquid culture) has been an ultimate goal in the field for *S. cerevisiae*. However, the absence of a strong transporter capable of enabling this trait has complicated the matter. Nonetheless, over the past few years, significant work complements the study here by identifying alternative mutations that influence sugar transport rate and selectivity.

Several of the mutational hot spots identified in this study parallel similar regions in *S. cerevisiae* transporters. For example, prior research has suggested the importance of an arginine residue on transmembrane domain (TMD) 8 and a tryptophan residue on TMD 5 for xylose uptake (Farwick *et al.* 2014b; Nijland *et al.* 2014; Shin *et al.* 2015). Here, we demonstrate that similar regions on the *C. intermedia* transporter are likewise essential. Thus, these key residues may be important harbingers of function when performing bioprospecting screens. However, the exact residue composition seems to be inconsistent across transporters. It was interesting to observe here that the N326H / T170N double mutant performed worse than the N326H single mutant yet the I171 mutation was additive with N326H. Likewise, the N326H mutant performed better than the phenylalanine mutation suggested in literature. Thus, it is likely that the specific amino acid composition in these sites will need to be optimized more globally with the transporter of interest. Certainly, a more relevant crystal structure model can help elucidate these differences further.

A unique finding in this work was the identification of a truncated C-terminal transporter tail which resulted in a decreased glucose inhibition. The current best model

for a yeast transporter protein comes from the crystal structure of the *E. coli* xylose transporter, XylE (Fedorova *et al.* 2012). The XylE transporter however, does not have a long C terminus tail like most yeast hexose/pentose transporters. Given the improvement we see in this work by truncating this tail, we hypothesize two potential functions for this tail region that can be further investigated in future work. First, it is possible that the C-terminal tail functions as an auto-inhibitory gate that typically responds to glucose. Thus, the removal of this tail allows for xylose to be transported at a higher rate. Second, it is possible that the lysine residues within the C-terminal tail can be ubiquinated thus leading to transporter degradation. In this mechanism, removing the tail allows for the enhancement of transporter half-life and/or membrane localization traits, thus resulting in improved cell growth and sugar transport. Moreover, the enhancements seen in glucose inhibition may be due to changing the next number of active transporters on the membrane surface. While this is an appealing regulatory hypothesis, we should note that residue 491 and 489 are both lysine residues, but truncations that terminates at A487 grew worse than truncations that terminates at K497. Nevertheless, since the C terminus tail of CiGXS1 does not align well with other commonly studied transporters, it is not clear how universal such truncations will be for reducing glucose inhibition. Certainly this study warrants future experiments focused on the function of the C-terminal tail of other major facilitator superfamily transporters. Regardless, the basic notion that XylE can transport xylose without glucose inhibition with a much shorter C-terminus tail (27 amino acids) compared to yeast hexose

transporters (55 amino acids for CiGXS1 WT) suggests that the length and composition of the C-terminal tail may play a role in substrate preference.

Finally, the importance of the G-G/F-XXX-G motif has recently been confirmed in other experiments using the Hxt7 transporter (Reider Apel *et al.* 2016). The third variable residue (F40 in CiGXS1 and F79 in Hxt7) are clearly related to xylose uptake. Young et al. first discovered the F40V mutation and found that F40M mutation would completely eliminate glucose transport. Thus, it is not surprising that we discovered a change in this residue in our final round of mutagenesis moving from methionine to valine. The resulting transporter CiGXS1 FIVFH497* is able to transport glucose at the WT level while transporting xylose at even higher rate. At 70 g/L glucose, xylose transport was still observed, a feat that has not been reported with any other yeast transporter to date. These characteristics demonstrate the potential of this transporter for co-utilization applications. Further investigations are certainly warranted (and ongoing) for importing this transporter into a more industrially adequate strain other than EBY.VW4000 that has undergone significant chromosomal rearrangement that hamper full potential in complex fermentations (Solis-Escalante *et al.* 2015). Nevertheless, the results depicted in this work show strong co-utilization properties that can be tested in other strain. Moreover, this result demonstrate that hexose transporters (in particular, symporters) belonging to the major facilitator superfamily can be rewired to be both xylose exclusive and glucose-xylose co-transporters without inhibition by either substrate. Thus, this work addresses a remaining ongoing challenge in the field of xylose utilization in yeast.

Chapter 6: Conclusions and Future Work

For years, when metabolic engineers attempt to produce biochemicals using microbial systems, glucose was the presumed carbon source. Other carbon sources like glycerol or ethanol are only used when they alter the metabolite balance and thus result in higher yield. However, the addition of those chemicals makes the production process unprofitable for industrial production. While it is widely known that xylose is the second most abundant sugar in lignocellulosic hydrolysate, xylose utilization is infrequently studied in unconventional yeasts other than simply importing a heterologous xylose pathway into these organisms. In this study, a combination of modern and classical methods was used to enable the non-conventional yeast *Yarrowia lipolytica* to utilize xylose to produce its natural product-lipid, and also a diverse portfolio of other molecules. This work can be used as a guideline for future researchers to follow for enabling an unconventional yeast to produce biochemicals from a non-native carbon source.

In chapter 2, we showed that heterologous xylose pathway complementation into a new organism does not always result in robust xylose utilization. In this case, it was necessary to utilize a starvation adaptation method to activate the xylose pathway genes. Subsequent genome sequencing found that this higher expression level is not due to any single mutation, but duplication of the pathway genes. This starvation adaptation process is replicable in different strain background. The resulting strain can produce up to 15g/L lipid in bioreactor setting. This was among one of the first studies published able to activate the xylose pathway in *Y. lipolytica*, and the first study to report the large scale genome rearrangement. This project opened up the possibility for the rest of the chapters, and the genome replication can also be an interesting separate project in the future.

Based on the strains obtained in chapter 2, we further tested targets that were previously studied in a more common yeast, *S. cerevisiae*. These results suggested that only xylulose kinase is the bottle neck for xylose utilization in *Y. lipolytica*. This result showed that metabolic regulation in different yeast are very different and targets found through model organisms are not universal. One potential future direction to continue improving xylose in this host is to use CRISPR or the piggyBac system to identify other potential targets which promotes xylose utilization that are unique to *Y. lipolytica*.

In addition to native lipids as a product, this work also showed that *Y. lipolytica* can use xylose to produce other chemicals like α-linolenic acid, riboflavin, and triacetic acid lactone through a yeast mating approach. The mating phenotype of *Y. lipolytica* has been well studied. But using mating to combine phenotypes of different strains has never been demonstrated in *Y. lipolytica*. These results demonstrate that the mating approach is astonishingly effective. The mated diploid strains showed a very high level of productivity and stability. Therefore, this mating approach should be considered more frequently as researchers try to expand carbon source for microorganisms, especially when a cellular adaptation step is required. Future work in this project area includes testing whether constructing strains without sporulation gene can create diploid strains with higher stability. Furthermore, a broader collection of xylose utilizing strains can be created to match with different biochemical production strains to achieve highest productivity. Put another way, not all applications will want to have a lipid producing strain as the mating partner.

Finally, in Chapter 5, we attempt to further improve xylose utilization by engineering membrane transporter proteins to enable glucose/xylose co-utilization. The mutant transporter showed preference of xylose with no glucose inhibition in a *S. cerevisiae* strain. This work demonstrated, for the first time, that a predominant glucose

transporter can be evolved to have its preference shift to xylose, even at that exclusion of glucose. This transporter needs to be further tested in *Y. lipolytica* (as well as other strains) as the next step to further improve sugar co-utilization by fungi.

Overall, these studies prove that xylose can serve as a major carbon source for the production of many biochemicals beyond ethanol. Furthermore, many of the approaches used in this work can be generalized to other organisms or carbon sources.

Chapter 7: Materials and methods

7.1 METHODS FOR CHAPTER 2

7.1.1 Strains and plasmids

A list of the plasmids and strains used in this study is provided in Table 7-1. *Escherichia coli* strain DH10β was used for cloning and plasmid propagation. The list of primers used in this study is shown in Table 7-2. E26 strain is an evolved *Y. lipolytica* strain that can efficiently transform glucose into lipid (Liu *et al.* 2015b). The same strain with recovered *LEU2* and *URA3* was used for integration. *XYL1* and *XYL2*, or *XylA* and *XYL3* were cloned into pUCS integration plasmids respectively(Blazeck *et al.* 2014). The sequences of *XYL1*, *XYL2*, *XylA* and *XYL3* are codon optimized for *Y. lipolytica*. In pUCS plasmids, the inserted genes were driven by a strong hybrid promoter and the whole expression cassettes were flanked by 26S ribosomal DNA regions, to target for higher integration efficiency (Blazeck *et al.* 2013b, 2014). Linearized integration plasmids with xylose utilization pathway were transformed into *Y. lipolytica* cells, and then selected on plates lacking leucine or uracil to confirm integration.

Table 7-1: List of plasmids and strains used in Chapter 2.

Plasmids/strains	Description
Plasmids	
pUCS1-16TEF-XYL1	Integration plasmid of XYL1with URA3
pUCS2-16TEF-XYL2	Integration plasmid of XYL2with LEU2
pUCS1-16TEF-XylA3	Integration plasmid of XylA3with URA3
pUCS2-16TEF-XYL3	Integration plasmid of XYL3with LEU2
pUC19-Actin-XYL1-XYL2-Tail	
Y. lipolytica strains	
PO1f	Control strain
E26	Base strain
E26 leu^- ura^-	E26 strain with LEU2 and URA3 knockout by Cre recombinase
E26 XYL12	E26 leu^- ura^- with XYL1 and XYL2 integrated into genome
E26 XUS	Starved E26 XYL12 strain with xylose utilization ability

Table 7-2: List of primers used in Chapter 2.

Genes	Primers	Sequences
SS Xyl1	Xyl1 fwd	AGGCGCGCCATGCCTTCTATTAAGTTG
	Xyl1 rev	TGCTTTATAGAGAGGAAATAATTAATT
SS Xyl2	Xyl2 fwd	AGGCGCGCCATGACTGCTAACCCTTCC
	Xyl2 rev	CTCATTGACGGCCCTGAGTAATTAATTA
SS Xyl3	Xyl3 fwd	AGGCGCGCCATGACCACTACCCCATTT
	Xyl3 rev	GAAAGTGAATTGAAACACTAATTAATT
SS XylA3	XylA fwd	AGGCGCGCCATGGCTAAAGAATATTTC
	XylA rev	ATTATCGCGATGTATCAATAATTAATTA
Tail PCR primers		
Xyl1	Xyl1 SP1 D	CCCATGGGACTGGGACAAGA
	Xyl1 SP2 D	TTCCTATCTTCGTCTAAGAA
	Xyl1 SP3 D	GGTTGCTTTATAGAGAGGAA
Xyl2	Xyl2 SP1 D	TACGACTTGGTCAGAGCCG
	Xyl2 SP2 D	TAAGGGTGCTGTCAAGTGTC
	Xyl2 SP3 D	TCATTGACGGCCCTGAGTAA
26S	26S D	CTTCCTATCATACCGAAGCA
RT-PCR primers		
Actin	Actin fwd	AGTCCAACCGAGAGAAGATG
	Actin rev	GGGGAGAGAGAAACCAGAG
Xyl1	Xyl1 fwd	CACCGCTTACTCTTCGTTC
	Xyl1 rev	AGCCTTGATAGTTTCGTTCTC
Xyl2	Xyl2 fwd	CTGCTGTTGGAATCTTTGAC
	Xyl2 rev	GCGTCCTTGAACTTGTATCTG
C chromosome	C chrom fwd	CAAAACACCACGGTAATCGG
	C chrom rev	ACAGCAACTACTCCTTTCAC

7.1.2 Medium and cell culturing

E. coli strains were cultured in Lysogeny broth (LB) media supplemented with 100 μg/mL ampicillin. For integrating xylose pathway genes, 6.7 g/L Yeast Nitrogen Base (YNB) w/ammonia sulfate w/o amino acids (Becton, Dickinson, and Company) and 0.79 g/L Complete Supplement Mixture (CSM) –Leu –Ura (MP Biomedicals) and 20 g/L glucose was used for selection. For genomic copy number and expression level characterization, YNB with complete CSM was formulated with 20 g/L glucose or xylose. For characterizing growth in 50 mL flasks, only YNB was used with the corresponding carbon source. YPD media (yeast extract, peptone and dextrose) was used for testing strain stability. YNB with 10 g/L ammonium sulfate and 160 g/L xylose was used in bioreactor fermentations.

E. coli cultures were grown at 37℃ with constant shaking at 225 rpm. Small volumes of *Y. lipolytica* were grown in 3 mL media at 30℃ with 225 rpm shaking. Medium volume cultures were grown at 50 mL volume in 125 mL flasks under same conditions.

7.1.3 Bioreactor fermentations

Minimal media with only YNB and xylose were used to run large scale bioreactor fermentations. Batches of 1.5 L fermentations in 2 L bioreactors were inoculated to initial OD_{600} at 0.1. The temperature was maintained at 28℃. and pH was maintained no lower than 3.5 with 1 M NaOH. The dissolved oxygen level was maintained at 50% air saturation by varying agitator speed between 250 rpm and 800 rpm with a constant air input flowrate of 3.75 L/min. 10 mL samples were taken every 24 hours; OD_{600}, sugar concentration and

lipid production were measured for each time point. The entire fermentations lasted 216 hours. The model of bioreactors we used was New Brunswick™ BioFlo®/CelliGen® 115.

7.1.4 Cell starvation

Y. lipolytica strains E26 or PO1f with integrated *XYL1*, *XYL2* or *XylA*, *XYL3* were grow up in 3 mL YNB + CSM+20 g/L glucose at 30℃ and 225 rpm in a culture tube for 14 days. Then cells were washed with water and then inoculated into 50 mL CSM + xylose (20 g/L). Cell culture of E26 *XYL1XYL2* that used up xylose was streaked out on YPD plates to select for single colonies. 20 colonies were selected then grown up in 3 mL YPD. The grown up cell culture were inoculated onto YNB + CSM + 20 g/L xylose plates to select for strains with stable ability to use xylose. Seven strains were picked and glycerol stocked for further characterization.

7.1.5 HPLC AND GC ANALYSIS

Glucose and xylose concentrations were measured with HPLC. Supernatant from 1 mL samples were filtered through 0.2 μm filters prior to running HPLC. A Thermo Scientific Dionex UltiMate 3000 Rapid Separation LC (RSLC) system, with RefractoMax 521 Refractive Index Detector and Aminex HPX-87P Column were used to measure samples. Filtered and degassed H2O was used as mobile phase. The flow rate of mobile phase was 0.6 mL/min. The column was kept at 85℃. Lipid samples were extracted and prepared following the procedures described in Liu L. et al (Liu *et al.* 2015a). TRACE™

1310 Gas Chromatograph was used for measuring lipid samples. Hexane was used as solvent for esterified fatty acids.

7.1.6 TAIL-PCR

To determine integration sites of *XYL1* and *XYL2*, TAIL-PCR was performed according to protocol previously described (Liu and Whittier 1995). Primers near 3' ends of *XYL1* and *XYL2* were paired with degenerate primers to amply 3' end of integration locus on genomic DNA. Three rounds of amplification were performed for each strain, and the amplified products were purified and sequenced by Sanger sequencing.

7.1.7 RT-PCR

RiboPure™ RNA Purification Kit (yeast) from ThermoFisher Scientific was used to extract total RNAs from *Y. lipolytica* samples. High-Capacity cDNA Reverse Transcription Kit from ThermoFisher Scientific was then used to convert mRNA into cDNA. Wizard® Genomic DNA Purification Kit from Promega was used for genomic DNA extraction. FastStart SYBR Green Master from Roche was used for RT-PCR. A plasmid constructed from pUC19 with actin (YALI0D08272g), *XYL1* and *XYL2* were used for generating standard reference curve. Copy numbers were calculated using the standard curve, and then *XYL1* and *XYL2* relative copy numbers of expression were plotted against copy numbers of actin driven by its native promoter.

7.1.8 Bioinformatics

Illumina HiSeq 2500 Run pair end sequencing PE 2X125 was used to sequence genomic DNA extracted from strain E26 XUS. The HiSeq generated 25,337,178 read pairs

that covered the *Y. lipolytica* genome approximately 225 times. The sequencing results were mapped to the CLIB122 genome using BWA; and Samtools, BEDTools and IGV were used to analyze the data (Li and Durbin 2009; Li *et al.* 2009; Quinlan and Hall 2010; Meacham *et al.* 2011).

CHAPTER 7.2 METHODS FOR CHAPTER 3

7.2.1 Strains and plasmids

Y. lipolytica strains used in this study are described in Table 7-3. *Escherichia coli* strain DH10β were used for cloning. Plasmids used in this study is provided in Table 7-4. PO1f XUS was described in prior publication (Li and Alper 2016) and briefly described here. PO1f Xylose Utilizing Strain (XUS) was generated by integrating Xyl1 and Xyl2 into the genome of the respective background strain followed by a starvation adaptation that resulted in increased copy number and expression (Li and Alper 2016). Primers and codon optimized spXYL1.2 was synthesized by IDT (Integrated DNA technologies).

Table 7-3. List of strains used in Chapter 3.

Strain name	Genotype
PO1f XUS	PO1f 4xXYL1 4xXYL2
PO1f XUS XKS	PO1f 4xXYL1 4xXYL2 ylXKS
PO1f XUS TAL	PO1f 4xXYL1 4xXYL2 ylTAL
PO1f XUS TKL	PO1f 4xXYL1 4xXYL2 ylTKL
PO1f XUS RKI	PO1f 4xXYL1 4xXYL2 ylRKI
PO1f XUS RPE	PO1f 4xXYL1 4xXYL2 ylRPE
PO1f XUS XKS TAL	PO1f 4xXYL1 4xXYL2 ylXKS ylTAL
PO1f XUS XKS TKL	PO1f 4xXYL1 4xXYL2 ylXKS ylTKL
PO1f XUS XKS TAL TKL	PO1f 4xXYL1 4xXYL2 ylXKS ylTAL ylTKL
PO1f XUS XKS YEF	PO1f 4xXYL1 4xXYL2 ylXKS ylYEF1
PO1j XYL1 XYL2 XKS	PO1j 1xXYL1 1xXYL2 ylXKS
PO1j XYL1.2 XYL2 XKS	PO1j 1xspXYL1.2 1xXYL2 ylXKS

Table 7-4. List of plasmids used in this Chapter 3.

Plasmid name	*Function*
pUC-Hph-ylXKS	ylXKS integration plasmid
pUC-Hph-ylTAL	ylTAL integration plasmid
pUC-Hph-ylTKL	ylTKL integration plasmid
pUC-Hph-ylRPE	ylRPE integration plasmid
pUC-Hph-ylRKI	ylRKI integration plasmid
pUC-NAT-ylTAL	ylTAL integration plasmid
pUC-NAT-ylTKL	ylTKL integration plasmid
pUC-GuaB-ylTKL	ylTKL integration plasmid
pUC-NAT-ylYEF1	ylYEF1 integration plasmid
pUCS3-spXYL1.2	spXYL1.2 integration plasmid
pUCS3-XYL1	XYL1 integration plasmid
pUCS2-XYL2	XYL2 integration plasmid

7.2.2 Media conditions

E. coli strains were cultured in Lysogeny broth (LB) media supplemented with 100 μg/mL ampicillin at 37°C. For cultivation of *Y. lipolytica*, YSC media with 40 g/L xylose, 0.79 g/L CSM-Complete (MP Biomedicals), and 6.7 g/L Yeast Nitrogen Base (Becton, Dickinson, and Company) was used for flask fermentation. *Y. lipolytica* flask fermentations were routinely performed at 28°C in 250 ml flasks with 225 rpm shaking. Yeast Peptone Glucose media (YPD) containing 10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose was used to prepare cells for transformation.

7.2.3 HPLC

Sugar concentrations were measured using a Thermo Scientific Dionex UltiMate 3000 Rapid Separation LC (RSLC) system with UV detector or RefractoMax 521 Refractive Index Detector as required. This measurement was likewise conducted as previously described without modification (Li and Alper 2016).

CHAPTER 7.3 METHODS FOR CHAPTER 4

7.3.1 Strains and plasmids

Y. lipolytica strains used in this study are described in Table 7-5. *Escherichia coli* strain DH10β were used for cloning. Plasmids used in this study is provided in Table 7-6. *Y. lipolytica* strains E26 XUS, ALA-A, RIB-A and TAL-A were described in prior publications and briefly described here (Li and Alper 2016; Cordova and Alper 2018; Markham *et al.* 2018; Wagner *et al.* 2018). E26 Xylose Utilizing Strain (XUS) were generated by integrating Xyl1 and Xyl2 into the genome of the respective background strain followed by a starvation adaptation that resulted in increased copy number and expression (Li and Alper 2016). ALA-A strain overproduces α-linolenic acid and was generated by integrating multiple copies of RkD12-15 into the background strain L36DGA1 (Cordova and Alper 2018). RIB-A strain overproduces riboflavin and was generated by overexpression of RIB1 and RIB3 followed by multiple rounds of mutagenesis and selection (Wagner *et al.* 2018). TAL-A strain overproduces triacetic acid lactone and was constructed by first integrating 4 copies of the polyketide synthase, g2ps1, followed by an overexpression of ACS1, ALD5, PDC2 and ACC1 (Markham *et al.* 2018). XUS-B mKate and PO1f-A hrGFP strains were generated in this study by integrating mKate2 or hrGFP with integration plasmid and overexpressing with a strong hybrid promoter.

Donor DNA for the mating region including MATB1 and MATB2 was synthesized by IDT using the sequence annotated for strain CLIB122. This synthesized sequence was then cloned alongside a NAT marker flanked by the piggyBac ITR

sequence. Finally, 1 kb of upstream and downstream homology sequence was cloned to match the mating region. The overall sequence for this construct is provided in Table 7-7. The entire integration cassette was PCR amplified into two sections using a homology overlap region in the NAT marker prior to transforming into the xylose utilizing mating type A strains via electroporation (Markham, Vazquez and Alper 2018). Putative transformants were selected on a YPX+NAT plate, and confirmed by PCR from outside the homologous region. A plasmid expressing excision-only version of piggyBac transposase was then transformed into this strain to remove NAT marker as per previously reported (Wagner, Williams and Alper 2018). Positive confirmation of the NAT removed strain was confirmed by PCR with primers outside of homologous region.

Table 7-5. List of strains used in Chapter 4.

STRAIN NAME	*GENOTYPE*
E26 XUS	E26 4xXYL1 4xXYL2 MATA1 MATA2
E26 XUS XKS	E26 4xXYL1 4xXYL2 ylXKS MATA1 MATA2
XUS-B	E26 4xXYL1 4xXYL2 ylXKS MATB1 MATB2
XUS-B MKATE	E26 4xXYL1 4xXYL2 ylXKS mKATE2 MATB1 MATB2
PO1F-A HRGFP	PO1f hrGFP MATA1 MATA2
ALA-A	L36DGA1 2xRkD12–15 MATA1 MATA2
RIB-A	PO1f RIB1 RIB3
TAL-A	PO1f 4xg2ps1 ACS1 ALD5 PDC2 ACC1

Table 7-6. List of plasmids used in Chapter 4

PLASMID NAME	*FUNCTION*
PUC-NAT-MKATE2	mKate2 integration plasmid
PUC-GUAB-HRGFP	huGFP integration plasmid
PUC-MATB CASSETTE PART I	Mating type switch plasmid
PUC-MATB CASSETTE PART II	Mating type switch plasmid
PMCS-HYEXC	ITR removing plasmid

Table 7-7. Sequence of MATB1 and MATB2 region with 1kb upstream and downstream sequence.

GACAGTGTAAGAGCTGGAGGGAAGCCAAAGTTGGCCGACGAGTCTTCCAC
AGACTGGTCCCCGACCTGTTACAGGACTCGTGCCGCAACAAGCATCCCAAT
CGACCCGACATGTATACTTGCTGGAACGTGATGATGAACTACCGGCCCGGA
AACTTTGGGTCTCGTATCGACTACGTGCTGGCCTCAAAAGGTCTGAAGGTG
GAGGATGCCGATATTCTGCCCCACCTGGAGGGCTCAGACCATTGTCCCGTT
TATGCAGAGGTGCGCGAACCAGACGCGTGCCAAGAAAAAGACAACCGTGC
CACTCAAGGTGCAGCTAACCAGACCAAACAAGCGTCGTCCTCAGCCCCAA
ACTCGTTTCGGCCAAAACTGTTCCACAGAACTAACAAGCAGTTTGTGCAGA
CCACCTCGGTCATGTCCATGTTCCGGGCGCAAACGCCGCCGACCAAGGGCA
TTACCAAACCCAAAGCGCCCTCCAAATCGACTGCCAAAAAACAGGCGGCC
ATCTCCTCCTTCTTCCAAAAGCCAGATCCAGAAGCTGCCGTCGTATCAACA
TGCGTCTCGGATGTGCACACAAAGCCGTTGGGACAATCACAATCACCCGCC
GTGTCCCCAAATATTGAAAGCACCTATCAAGGTTGCGATGCAGGTGCGCAT
GCGACTCGAGAGGGTGAAGATACGAGACCGCAAATGGCCACCAGACTGGA
CTTGCCAGAGCCCACCGAGTACCTAGACATTGCGGAGCGCGTGGAGGAGG
CCGAAAAGGTCGCGTCGCAGTGGTCCAAGATCATGACAAAAAAACCAGCC
CCCGTCTGTGCCCACAACCTCCCGGCGGTGCTAAGAACCACCCGCAAAAAG
GGACCCAACCAGGGCAAGCACTTTTGGGTTTGTGGCCAGAACCACTCGCTG
CAGCTCGTCGGCGACTCCTACAAGGCATCGGCAGACCCCGACGCCAAGGG
CTCCCAGGACGACACCAGTGCCAGATGTAACTTCTTTCAATGGAAATGATC
GAAGGTTGGCTATTGTAACAAGGAAACTGCCATTTCAGATGTTTACAACTG
CTTTTGATATATTATACTTTTATAAAAGATGAGGATCGACAACGATAATGC
ACAATGAGCCTGTTTCACCAACACTGTCATCCCATTTGGGTTAAGTCAATAT
CAGCGTTCCTATTACCTAAAACCGAAGGGTCTACTACTTATATAGTATTGGT
AGTTTCTTTGACACGTCCACTCTCGTCTTACATATCATACCACTACCGAATG
AAGGAAACCAGAGATCTTCTCTTCTACAACGAGAACTCCATCGACTTCAAC
TTCTCTCCAGAAGAGATGGACACCCATCAAATGATGTGTGTCAACAATGAC
AGCCAACAGGACACATCGTTCCTCACAGCTACTACTTGCATTTTTGACTTCC
AGAAGGACATCTATGACGACCAAAACCTGATGTCCTTCGACCCCTTCTTTC
CACAGCAGCGAGATATGGCTTTGTGGATTGCTTCAGTGGGAAATGGGGCTA
TCGATTGGAGTACTAAACCTCAACTGACTCTTTCAAAGAAGATCCCTGACG
TCGTCAACGAAAAAGGAGAGGAGATAACCCCTGCTGACCTACAACGCCAC
TATGCCGATAAGCATGTCGCCAAACTTCCAGATCATTCCTGCGATCTTTACT
ACCTTCTATTCAAGCATGCCAAATTCGAGGCTACCACCCTGAGGATCCCAT
TGTCACTCCTCAATGACAACGGTACAGGAGCACTTGGCCCATTCGACGAGA
TCAGATCGTTGACGGATGAAGAACAACGTCTTCTTACTCTGCTTAAGACAA
TTTCACATTACAGCTGGGTCACTAAACATGCGACCAAAGACTGTACGGGCC
AACTGCTTGACCTGTGCAAAGAGGGCCTCCAGCTTGTGGAATCTATCCAGG
TTAACAAAACTTTCAACGGTGTCATTCTGCATATCCTAGTCGCCTACTTTGG
GAGAAGATCTACTGCTCTAAGAGATGAGTTTCAACTGCACAAGCTCTTTCT

Table 7-7 continued. Sequence of MATB1 and MATB2 region with 1kb upstream and downstream sequence.

TGAACAACGGCTTGACATTGAGTTGGAAATCTTGGCCCAACATGGAGGCGC
TGACGCCATTGATTCTGAAAAGCTGCGGAAGGTCAAGAAACTGAGCGAAC
TGTGGGAGTGGTTCAATGCCATCCCGTACAACCACCACCCTAGCGTCTCTC
AGATCGAGTTTATCGCTTCTCAGATCGGGGAGAAAGTAAGTTTCGTGAAAT
CCTGGGTTGCAAACAAGCGACGAACTGGGGCCAAGAAGAGATCCAAGAAA
CCGGCACCTTCGACACAGATTAGACCTGCATTGAAGGTATTCTTGGAGAAG
AAGCCATTTGTGTCTCGCCAAGTTGTGTAATTGAATCATAATGTATTACAAT
GTTGTTAGGTACATACTGTAGATTGTTTTATATGCAATATTGTGCCCAATAC
ACAAGGTACAGTATGTACAATACAATACCTAATGTTCATACTGAACACCTT
ATATACAAGTACATTAAGGTAGTTGTATGTATGTACATACTGTAGTCACTA
CAGACGCTATCTTTCCAGTGTGCCGACCGGTAACTGAAAGCAAATATAGAA
GTACGGTTTAAGAAAGATATCATTTAATCAAAGGTGCGATATCTAGAGGCT
TGCCCCATTCTTTGTTTGTGCAATGAACAGGGAACATGCTATGCAGTAACA
CCAAATAATGTACTTAGTCTAATCACCTTCTGTTATATTTAGTTATCTTCGC
TCCTCGTCTTATTGTCGTCAATCTAAGTAGGTACGGTCGCCGCAAAAGGAA
ATCTCTTGCCGCCGTCGCTACAGTGATCAAGTAGCTGAGCTACGAAGGTAA
TTAAAATATTTGTTCGCTTTGCTCATACGTCGTTTTAATCTTAAATTTCATCT
GATTACACTTCGGCGAGGAGTTGTTTCTGGCATCTCGCCCTTCGACCTTGCT
TATGATTTCTGTGACCAGAAACCTAAATTACCCACAATGTGAGTTGCAAAT
GCAGGCAGCTAATTACAGTACAGTACAGTACAGTACTGAACTGTATGTACA
GTAGGTCACGAAGAAGTTAATACTTGTGAGGCAGAGTACCACTTCACAGG
GATGCAATTGCAATGTAGAAAAATACTGTGAAAGTAAAGTATTAAACCTAT
CCAAGCAAAATAAAAGTGTTGACTCAAATAATTCAACCCTTTCCTACCAGC
AAGCAATGGAGAAGTAAAGACGAGGGACTGCCACCCATCCTCAAGCTGTT
ACTCAAGGTTTGAAAAAGAGTTCGGACACTCAACTTCCCCAGCCAAAATCG
CATAAACAAAGAATCTCGCGAGATTGGGTTTATCACCTAATGAGGTAATCC
TCACGATGAGGCAAGATGAGTATCATATAAAGCAAAGGATTTTTTCAATCT
CAAAAGCAGCTATAACAATGATCAAGATAAAAACCGAGCACAAAGTGGTG
AAGGCAAGAACCCCGAAAAAAATGAAACCTGGAATGGCCCGTTTCAGAAT
TCAAGCCGTTCCCATCATCAACCAAAACAACAACGCCTACCTGTCAGAGAT
CAGCTTCGAAGAAAAGTTAGAACCAGTTCCGCCGATGTATCCTGAACTACA
AAAAGTCCTTGATTACATTCACTCACGCAAAGTCCCCTTCGACGGATCATC
GGGGCCATACCCTCCCTCGAGCAGACGACAAGTGGTCAACGGTTATGTCGC
ATTTAAGCTATATCACCACGTTCCACACACTGACCTTAGCGCAATTGATATC
ACCCACCTCCTCCAAGGTCCTTGGAGAGCTTGCCCTGATAGACACATCTGG
ACCAAGTACGCTGATCACTACCGTTCTAGAGGGCGTGACGTCGCATTCAAG
ACATGGCTTCTGTCTGTTACTAGTCCTGCTCCTGAGAAGGGGCTAATTATTC
AGCCTCAAACTCAGAGCGAGCAGGGGTACAACTTCTCATCTTCCTCTGATT
TTTCAAGTTCTCCAGAACACAGTGAAACACTCCCTGTGTCATCTGGAGAGC
CCAGTACACCTGTTGACCCCTTCCTGATCGAGTTTGATCAAGGCATCGCGG

Table 7-7 continued. Sequence of MATB1 and MATB2 region with 1kb upstream and downstream sequence.

AGAAAGAGGCCAGCTTCACAAGTCCTTGCAACCAAAACTATGACTTGGATT
TCAGCCAGCTTCAGCCTCTAAGTATTCGCGTCCCTTATTTTAACCAGTACGC
TTCTTCCAGTAACTCTGGATTGGATGTTGACTACTTTGACGGTAGCACCCTA
GACACTCCTAGCTACATCGACACTTTCATCACACCATTGCACTAGTCATAAT
GGTGTAGACATGGGTTTGTACGAGTACTACGATACGACCTCAAACTCCAAC
GGTAGTCCTATTTATTCACATCGTTAATATCATTTCATTCATTTACCATACA
TGTTATACATACATCAGCCGCGCGTAGCTTAGTCCTCCTGGCCAGGAATGG
ACTCCTCCTCGTCCTGATGCAGGTAGGCGACCTTTCGCATCTGTCCAAGTCG
CTTTCGGGCGCCCTGCAGATCGTTTTCGAGCTTGAGAATTTCGACCTGCTGT
TCCATTTCGGTCGTCTTGAACTCGTGCTTGGAGAGGGCAGCGTAGTCTACCT
TTTCGTCCAGCTCGCCAAACTTCTTGTTGAGAATTTCCTGCACCTGCTTGAC
GAGGTTTCGGCAGGCGGAAGTAACCACCTTTGAACACTCTTCCAGCTTGTC
CTGGGTCTTGGACATGAAGGTAGCCTTTACTCGAGAGGCAGCAACCAGCTG
AGCAGTGGAAGCAGCGACGTTGTTGGAAGCCACAATGAGCTCCTCGAGAC
CAGAGGTTCGTCGGAGAGTGCCGTCGGCCTTCTCGATGAGAATGTTGGTCG
ATGCAGCTACACTCTTGGCCGCGGAAATAAGACCCTCGGTCCATTTGTTGT
TCTTCTTGTAGAACTGTGCTCGAGAAGAAGTACCTCGTCCCTGGGCTACAA
TCTCGTTCTGGGCATCGGTGGCAGCTCGAATCAGCGCAGCAATGGCGTTGA
TGACGGCCTGGGCAGCCTGGATGGCAGCCTCGTGCAGCTGCAGGTCGGTGG
TCGACTTGAGAGGGTCGGACTGGTTAAACTTGAGCAGATCGCCAAGCTTAG
CAGAAGCAGCAGCAATGGCATCAGCAGCCTTGGCCATCTCTCGCTCAACGA
GGTCGCCCAGATCACCAGACATGTTGATGTTGGTCTTGGGGGCCATTCGCT
CGGCCAGCTTGGTGAGGTGCTGAAGAGCAGTCTGCACGTTCATGTTGAGAG
TGATGACCCGCTCCATCTTGTCATCCAGAGGGAAATCCTCAATGATCTCAG
ACAGAAGACCAATAAAGTTTCGCTCTGTGGCCTCGGCAGACTCACGAGCAA
A

7.3.2 Media conditions

E. coli strains were cultured in Lysogeny broth (LB) media supplemented with 100 μg/mL ampicillin at 37°C. For cultivation of *Y. lipolytica*, YSC media with 40 g/L glucose or xylose, 0.79 g/L CSM-Complete (MP Biomedicals), and 6.7 g/L Yeast Nitrogen Base (Becton, Dickinson, and Company) was used for flask fermentation. *Y. lipolytica* flask fermentations were routinely performed at 28°C in 250 ml flasks with 225

rpm shaking. The ALA flask experiment was performed at 20 °C. Yeast Peptone Glucose or Xylose media (YPD or YPX) containing 10 g/L yeast extract, 20 g/L peptone and 20 g/L glucose or xylose was used to prepare cells for transformation or mating. YPX with double amount of xylose was used for ALA flask experiment. Yeast extract – malt medium (YM) consisting of 5 g/L peptone, 3 g/L yeast extract and 3 g/L malt extract (Austin Homebrew Supply) was used for mating.

7.3.3 Mating method

The mating method utilized here was adapted from a previously reported protocol (Barth & Gaillardin, 1996) with modifications. Briefly, parental strains were grown in 2 mL YPD media at 28 °C in tubes overnight, then 1 mL of each culture was combined into a fresh tube. The mixed cells were rinsed with YM medium and then resuspended in 10 mL YM medium. This mixture was incubated at 28°C without shaking for 1-4 days depending on strains used. Mating efficiency varies due do different parental strains used. The mated culture from different dates was plated on YPX with 1 mg/ml nourseothricin for selection. Isolated colonies were then purified by re-streaking onto fresh plates. The ploidy was confirmed using internal primers embedded in matA1 matA2 and matB1 matB2 genes.

To measure conjugation frequency, XUS-B mKate and PO1f-A hrGFP were mated according to the protocol described above. Both parental strains as well as the mated culture were enriched in YPD media for overexpression of the fluorescent proteins. The enriched cultures were run on flow cytometry (BD Accuri™ C6 Plus flow

cytometer) as well as the original cell mix before conjugation event occur. The parental strains were used for compensation and the original cell mix was used as reference to calibrate the growth rate difference. The diploid cells are the population that have fluorescence of both fluorescent proteins.

7.3.4 Bioreactor fermentation

We performed bioreactor fermentation as previously described (Cordova and Alper 2018), however, xylose was used in place of glucose (with the same mass concentrations used for the fermentation). Initial media contained 0.79 g/L CSM-Complete, 3.4 g/L Yeast Nitrogen Base and 80 g/L xylose. The bioreactor was inoculated with an initial concentration of 0.1 OD600 unit. As additional 80 g of xylose was added after the 72-hour timepoint to be consistent with the previous, glucose ALA bioreactor fermentations conditions. The total volume in the reactor was 1.7 L after the xylose addition. The pH of the bioreactor was single-sided control with 1 M NaOH to keep the value greater than 3.5 and dissolved oxygen was controlled to be no less than 50% saturation using an agitator cascade (maximum 800 rpm, minimum 250 rpm). The model of bioreactors used in this study was New Brunswick™ BioFlo®/CelliGen® 115.

7.3.5 Product analysis

Lipid content including ALA was extracted and analyzed using GC mass spectrometry (TRACE™ 1310 Gas Chromatograph) as previously described without modification (Cordova and Alper 2018). Riboflavin was measured at 450 nm wavelength using a spectrophotometer as described by Wagner *et al.* without modification (Wagner

et al. 2018). TAL and sugar concentrations were measured using a Thermo Scientific Dionex UltiMate 3000 Rapid Separation LC (RSLC) system with UV detector or RefractoMax 521 Refractive Index Detector as required. TAL was measured using C18 reverse flow column and glucose or xylose were measured using HPX-87P Column. These measurements were likewise conducted as previously described without modification (Li and Alper 2016; Markham *et al.* 2018).

CHAPTER 7.4 METHODS FOR CHAPTER 5

7.4.1 Strains, plasmids and media

S. cerevisiae EBY.VW4000 strain was provided by Dr. Eckhard Boles of the Institute of Molecular Biosciences, Goethe-Universität, at Frankfurt, Germany. Strain ETKX was constructed by deleting HXK1, HXK2 and GLK1 in EBY.VW4000. Plasmids p416GPD-X1X2X3 and p415Gal1p-Gal2 were transformed into ETKX strain to build strain ETKXG. ETKXG was able to grow on galactose and xylose but not glucose, and was used for testing and selecting xylose transporters. CiGXS1 transporters were subcloned into plasmid p414TEF. Quick Change PCR primers (Table 7-8) were used for making single mutations. CiGXS1 fwd primer were paired with C-terminus truncation primers (Table 7-8) to make seven transporters with different C-terminus tail lengths. The xylose utilization pathway was expressed in EBY.VW400 on plasmid p416GPD-X1X2X3. The resulting strain EBYX was used for characterizing the glucose and xylose co-utilization. In the co-utilization test, WT, FIM and FIVFH497* were expressed using the GPD promoter in plasmid p414GPD. All yeast shuttle vectors were previously described by Mumberg et al. (1995). Yeast strains were grown on Yeast Nitrogen Base (YNB) with ammonia sulfate without amino acids (Becton, Dickinson and Company. Sparks, MD USA), Complete Supplement Mixture (CSM) with certain dropout to maintain plasmids with selecting carbon sources. *Escherichia coli* strain DH10β (New England BioLabs, Ipswich, MA) was used for cloning and plasmid propagation and were cultured in Lysogeny broth (LB) media containing 100 μg/mL ampicillin.

Table 7-8. List of primers used in Chapter 5.

Mutagenesis primers		***Sequences***
	CiGXS1 fwd	GATTCTAGAACTAGTGGATCCATG
	CiGXS1 rev	GGTCGACGGTATCGATTTA
Quick Change PCR primers		
T170N	C509A fwd	GGAATAGTAGAACCGACATGGACAATGTTAGTGGCCAAG
	C509A rev	CTTGGCCACTAACATTGTCCATGTCGGTTCTACTATTCC
N326H	A976C fwd	GCCAAGAAAAGACCAATGTTAATAGCCCACTGGTAACAA
	A976C rev	TTGTTACCAGTGGGCTATTAACATTGGTCTTTTCTTGGC
N326F	A976T A977T fwd	GGAATAGTAGAACCGACAAAGACAATGTTAGTGGCCAAG
	A976T A977T rev	CTTGGCCACTAACATTGTCTTTGTCGGTTCTACTATTCC
M40V	A118G fwd	TGCTTCTGGTGGTTTCATTGTGGGATACGATACTGGTAC
	A118G rev	GTACCAGTATCGTATCCCACAATGAAACCACCAGAAGCA
C-terminus truncation primers		
	Y467* rev	AAAAAAATCGATCTAGATGAAGTACCAAGCGAAGA
	V477* rev	AAAAAAAATCGATCTACTGCTCCAAAGAAAGACCC
	A487* rev	AAAAAAAATCGATCTACTTGCTGACATGCTCGTAC
	K497* rev	AAAAAAAAAATCGATTTAAGATGGAACGAAGCCCT
	E502* rev	AAAAAATCGATCTACTCTCTGAAAGAGTGCTTAGATG
	Q507* rev	AAAAAAAAAATCGATTTACTGGTCCACCTGCTCTC

7.4.2 Directed evolution

Random mutagenesis of *C. intermedia* GXS1 using mutagenesis primers (Table 7-8) was accomplished with Gene-morph Mutazyme II kit (Agilent Technologies, Santa Clara, CA). Mutagenesis libraries were constructed as previously described (Young *et al.* 2012). These libraries were then transformed into *S. cerevisiae* ETKXG. The first three rounds of libraries were plated on CSM-Ura Leu Trp minimal medium with 20 g/L xylose and 20/40/60 g/L glucose plates. Large colonies were then tested again in 3 ml of the same media in culture tubes in parallel with their parental strain to confirm for better growth. In the fourth round of mutagenesis, the yeast strain transformed with transporter library was grown in 50 mL shake flask with CSM-Ura Leu Trp minimal medium with 20 g/L xylose and 60 g/L glucose. The combined culture went under three rounds of dilution before plating on plates for single colonies. Strains with confirmed better growth were sequenced to check for mutations.

7.4.3 Growth condition

E. coli cultures were grown at 37℃ with constant shaking at 225 rpm. *S. cerevisiae* were grown in 3 mL media in 14 mL culture tube or 50 mL in 125 mL flask at 30℃ with 225 rpm shaking. For deep well assay, yeast strains were grown in 1.4 mL media in 96 well deep-well plates with 1.8 mL capacity for each well. The deep-well plates were incubated at 30℃ with 1000 rpm shaking. 100 μL samples were taken every 24 hours and relative OD_{600} was measured in clear bottom 96 well plates using plate reader.

7.4.4 High Cell Density Flask fermentation

Strain EBYX transformed with p414TEF-CiGXS1 wild type transporter, FIM and FIVFH497* mutants was first grown in CSM-Ura Leu with 20 g/L xylose as start culture. Cells were then washed and inoculated into 50 mL of CSM-Ura Leu with 10 g/L xylose and 10 g/L glucose. The initial OD for the flask fermentation was 3 for each flask. Sample were taken every 24 hours to measure OD and sugar content.

7.4.5 HPLC analysis of sugars

Glucose and xylose concentrations were measured using HPLC. Supernatant from 1 mL samples were filtered through 0.2 μm filters prior to running HPLC. A Thermo Scientific Dionex UltiMate 3000 Rapid Separation LC (RSLC) system (Sunnyvale, CA USA), with RefractoMax 521 Refractive Index Detector (Kawaguchi, Saitama Japan), and Aminex HPX-87P Column (Hercules, CA USA) were used to measure samples. Filtered and degassed H2O was used as mobile phase, with a flow rate of 0.6 mL/min. The column was kept at 85℃ during the measurement.

References

Abdel-Mawgoud AM, Markham KA, Palmer CM *et al.* Metabolic engineering in the host *Yarrowia lipolytica*. *Metab Eng* 2018, DOI: 10.1016/j.ymben.2018.07.016.

Abu-Hamdeh NH, Alnefaie KA. A comparative study of almond and palm oils as two bio-diesel fuels for diesel engine in terms of emissions and performance. *Fuel* 2015, DOI: 10.1016/j.fuel.2015.02.040.

Adesanya VO, Cadena E, Scott SA *et al.* Life cycle assessment on microalgal biodiesel production using a hybrid cultivation system. *Bioresour Technol* 2014, DOI: 10.1016/j.biortech.2014.04.051.

Amiri H, Karimi K. Autohydrolysis: A promising pretreatment for the improvement of acetone, butanol, and ethanol production from woody materials. *Chem Eng Sci* 2015, DOI: 10.1016/j.ces.2015.07.020.

Ammar EM, Wang X, Rao C V. Regulation of metabolism in *Escherichia coli* during growth on mixtures of the non-glucose sugars: Arabinose, lactose, and xylose. *Sci Rep* 2018, DOI: 10.1038/s41598-017-18704-0.

Anderlund M, Rådström P, Hahn-Hägerdal B. Expression of bifunctional enzymes with xylose reductase and xylitol dehydrogenase activity in *Saccharomyces cerevisiae* alters product formation during xylose fermentation. *Metab Eng* 2001;**3**:226–35.

Balan V, Bals B, Chundawat SPS *et al.* Lignocellulosic biomass pretreatment using AFEX. 2009.

Bamba T, Hasunuma T, Kondo A. Disruption of PHO13 improves ethanol production via the xylose isomerase pathway. *AMB Express* 2016;**6**:4.

Barth G. Mating, Sporulation and Random Spore Selection by Nystatin in *Yarrowia lipolytica*. *Non-Conventional Yeasts Genet Biochem Biotechnol* 2011:335–8.

Barth G, Gaillardin C. Physiology and genetics of the dimorphic fungus. *FEMS Microbiol Rev* 1997;**19**:219–37.

Barth G, Weber H. Improvement of sporulation in the yeast *Yarrowia lipolytica*. *Antonie Van Leeuwenhoek* 1985;**51**:167–77.

Beopoulos A, Mrozova Z, Thevenieau F *et al.* Control of lipid accumulation in the yeast

Yarrowia lipolytica. *Appl Environ Microbiol* 2008, DOI: 10.1128/AEM.01412-08.

Blazeck J, Hill A, Jamoussi M *et al.* Metabolic engineering of *Yarrowia lipolytica* for itaconic acid production. *Metab Eng* 2015;**32**:66–73.

Blazeck J, Hill A, Liu L *et al.* Harnessing *Yarrowia lipolytica* lipogenesis to create a platform for lipid and biofuel production. *Nat Commun* 2014;**5**:3131.

Blazeck J, Liu L, Knight R *et al.* Heterologous production of pentane in the oleaginous yeast *Yarrowia lipolytica*. *J Biotechnol* 2013a;**165**:184–94.

Blazeck J, Liu L, Redden H *et al.* Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. *Appl Environ Microbiol* 2011;**77**:7905–14.

Blazeck J, Reed B, Garg R *et al.* Generalizing a hybrid synthetic promoter approach in *Yarrowia lipolytica*. *Appl Microbiol Biotechnol* 2013b;**97**:3037–52.

Boe A, Lee DK. Genetic variation for biomass production in prairie cordgrass and switchgrass. *Crop Sci* 2007, DOI: 10.2135/cropsci2006.05.0323.

Bozan B, Temelli F. Chemical composition and oxidative stability of flax, safflower and poppy seed and seed oils. *Bioresour Technol* 2008, DOI: 10.1016/j.biortech.2007.12.009.

Brethauer S, Studer MH. Biochemical conversion processes of lignocellulosic biomass to fuels and chemicals – A Review. *Chim Int J Chem* 2015;**69**:572–81.

Bruinenberg PM, de Bot PHM, van Dijken JP *et al.* The role of redox balances in the anaerobic fermentation of xylose by yeasts. *Eur J Appl Microbiol Biotechnol* 1983, DOI: 10.1007/BF00500493.

Cadete RM, De Las Heras AM, Sandström AG *et al.* Exploring xylose metabolism in *Spathaspora* species: XYL1.2 from *Spathaspora passalidarum* as the key for efficient anaerobic xylose fermentation in metabolic engineered *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2016;**9**:1–14.

Carlson R. Laying the foundations for a bio-economy. *Syst Synth Biol* 2007, DOI: 10.1007/s11693-007-9010-z.

Chiaramonti D, Prussi M, Ferrero S *et al.* Review of pretreatment processes for lignocellulosic ethanol production, and development of an innovative method.

Biomass and Bioenergy 2012;**46**:25–35.

Clifton-Brown J, Harfouche A, Casler MD *et al.* Breeding progress and preparedness for mass-scale deployment of perennial lignocellulosic biomass crops switchgrass, miscanthus, willow and poplar. *GCB Bioenergy* 2019, DOI: 10.1111/gcbb.12566.

Cordova LT, Alper HS. Production of α-linolenic acid in *Yarrowia lipolytica* using low-temperature fermentation. *Appl Microbiol Biotechnol* 2018;**102**:8809–16.

Czajka JJ, Nathenson JA, Benites VT *et al.* Engineering the oleaginous yeast *Yarrowia lipolytica* to produce the aroma compound β-ionone. *Microb Cell Fact* 2018, DOI: 10.1186/s12934-018-0984-x.

Dale BE, Ong RG. Energy, wealth, and human development: Why and how biomass pretreatment research must improve. *Biotechnol Prog* 2012;**28**:893–8.

David Martin Alonso JQB and JAD. Catalytic conversion of biomass to biofuels. *Biomass Conv Bioref* 2011;**1**:203–15.

Demeke MM, Dietz H, Li Y *et al.* Development of a D-xylose fermenting and inhibitor tolerant industrial *Saccharomyces cerevisiae* strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol Biofuels* 2013;**6**:89.

Donato P Di, Poli A, Taurisano V *et al.* Polysaccharides from bioagro-waste for new biomolecules. *Polysaccharides: Bioactivity and Biotechnology*. 2015.

Dulermo T, Lazar Z, Dulermo R *et al.* Analysis of ATP-citrate lyase and malic enzyme mutants of *Yarrowia lipolytica* points out the importance of mannitol metabolism in fatty acid synthesis. *Biochim Biophys Acta - Mol Cell Biol Lipids* 2015;**1851**:1107–17.

Farwick A, Bruder S, Schadeweg V *et al.* Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc Natl Acad Sci U S A* 2014a;**111**:5159–64.

Farwick A, Bruder S, Schadeweg V *et al.* Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc Natl Acad Sci U S A* 2014b;**111**:5159–64.

Fedorova T V., Chulkin A. M., Vavilova E. A. *et al.* Purification, biochemical characterization, and structure of recombinant endo-1,4-β-xylanase XylE. *Biochem* 2012;**77**:1190–8.

Frei M. Lignin: Characterization of a multifaceted crop component. *Sci World J* 2013, DOI: 10.1155/2013/436517.

Gaillardin CM, Charoy V, Heslot H. A study of copulation, sporulation and meiotic segregation in *Candida lipolytica*. *Arch Mikrobiol* 1973, DOI: 10.1007/BF00409513.

Galitski T, Saldanha AJ, Styles CA *et al.* Ploidy regulation of gene expression. *Science (80-)* 1999, DOI: 10.1126/science.285.5425.251.

Gibson B, Geertman JMA, Hittinger CT *et al.* New yeasts-new brews: Modern approaches to brewing yeast design and development. *FEMS Yeast Res* 2017;**17**:1–13.

Gírio FM, Fonseca C, Carvalheiro F *et al.* Hemicelluloses for fuel ethanol: A review. *Bioresour Technol* 2010;**101**:4775–800.

Gonçalves FAG, Colen G, Takahashi JA. *Yarrowia lipolytica* and its multiple applications in the biotechnological industry. *Sci World J* 2014;**2014**, DOI: 10.1155/2014/476207.

Harner NK, Wen X, Bajwa PK *et al.* Genetic improvement of native xylose-fermenting yeasts for ethanol production. *J Ind Microbiol Biotechnol* 2015;**42**:1–20.

Hilpmann G, Becher N, Pahner FA *et al.* Acid hydrolysis of xylan. *Catal Today* 2016;**259**:376–80.

Iranmahboob J, Nadim F, Monemi S. Optimizing acid-hydrolysis: A critical step for production of ethanol from mixed wood chips. *Biomass and Bioenergy* 2002, DOI: 10.1016/S0961-9534(02)00016-8.

Jagtap SS, Rao C V. Microbial conversion of xylose into useful bioproducts. *Appl Microbiol Biotechnol* 2018, DOI: 10.1007/s00253-018-9294-9.

Jeffries TW. Utilization of xylose by bacteria, yeasts, and fungi. *Adv Biochem Eng Biotechnol* 1983;**27**:1–32.

Jeffries TW, Jin YS. Metabolic engineering for improved fermentation of pentoses by yeasts. *Appl Microbiol Biotechnol* 2004;**63**:495–509.

Jin M, Bothfeld W, Austin S *et al.* Effect of storage conditions on the stability and fermentability of enzymatic lignocellulosic hydrolysate. *Bioresour Technol* 2013;**147**:212–20.

Jin M, Slininger PJ, Dien BS *et al.* Microbial lipid-based lignocellulosic biorefinery: feasibility and challenges. *Trends Biotechnol* 2015a, DOI: 10.1016/j.tibtech.2014.11.005.

Jin M, Slininger PJ, Dien BS *et al.* Microbial lipid-based lignocellulosic biorefinery: Feasibility and challenges. *Trends Biotechnol* 2015b;**33**:43–54.

Jin Y, Alper H. Improvement of xylose uptake and ethanol production in recombinant *Saccharomyces cerevisiae* through an inverse metabolic engineering approach. *Appl Environ Microbiol* 2005;**71**:8249–56.

Jones GA, Warner KJ. The 21st century population-energy-climate nexus. *Energy Policy* 2016;**93**:206–12.

Karhumaa K, Hahn-Hägerdal B, Gorwa-Grauslund MF. Investigation of limiting metabolic steps in the utilization of xylose by recombinant *Saccharomyces cerevisiae* using metabolic engineering. *Yeast* 2005;**22**:359–68.

Kim B, Du J, Eriksen DT *et al.* Combinatorial design of a highly efficient xylose-utilizing pathway in *Saccharomyces cerevisiae* for the production of cellulosic biofuels. *Appl Environ Microbiol* 2013;**79**:931–41.

Klinke HB, Thomsen AB, Ahring BK. Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl Microbiol Biotechnol* 2004;**66**:10–26.

Kumar P, Barrett DM, Delwiche MJ *et al.* Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Ind Eng Chem Res* 2009;**48**:3713–29.

Kurisciiko C. Parasexual process in the yeast *Yarrowia lipolytica*. 1986;**26**:33–41.

Kuyper M, Harhangi HR, Stave AK *et al.* High-level functional expression of a fungal

xylose isomerase: The key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res* 2003;**4**:69–78.

Kwak S, Jo JH, Yun EJ *et al.* Production of biofuels and chemicals from xylose using native and engineered yeast strains. *Biotechnol Adv* 2019, DOI: 10.1016/j.biotechadv.2018.12.003.

Larroude M, Rossignol T, Nicaud JM *et al.* Synthetic biology tools for engineering *Yarrowia lipolytica*. *Biotechnol Adv* 2018;**36**:2150–64.

Leandro MJ, Gonçalves P, Spencer-Martins I. Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose-H+ symporter. *Biochem J* 2006;**395**:543–9.

Ledesma-Amaro R, Lazar Z, Rakicka M *et al.* Metabolic engineering of *Yarrowia lipolytica* to produce chemicals and fuels from xylose. *Metab Eng* 2016;**38**:115–24.

Lee S-M, Jellison T, Alper HS. Systematic and evolutionary engineering of a xylose isomerase-based pathway in. *Biotechnol Biofuels* 2014;**7**:122.

Lee SM, Jellison T, Alper HS. Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2012;**78**:5708–16.

Li H, Alper HS. Enabling xylose utilization in *Yarrowia lipolytica* for lipid production. *Biotechnol J* 2016;**11**:1230–40.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;**25**:1754–60.

Li H, Handsaker B, Wysoker A *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;**25**:2078–9.

Lim YP, Go MK, Yew WS. Exploiting the biosynthetic potential of type III polyketide synthases. *Molecules* 2016, DOI: 10.3390/molecules21060806.

Liu L, Alper HS. Draft Genome Sequence of the Oleaginous Yeast *Yarrowia lipolytica* PO1f, a Commonly Used Metabolic Engineering Host. *Genome Announc* 2014;**2**:e00652-14.

Liu L, Markham K, Blazeck J *et al.* Surveying the lipogenesis landscape in *Yarrowia*

lipolytica through understanding the function of a Mga2p regulatory protein mutant. *Metab Eng* 2015a;**31**:102–11.

Liu L, Pan A, Spofford C *et al.* An evolutionary metabolic engineering approach for enhancing lipogenesis in *Yarrowia lipolytica*. *Metab Eng* 2015b;**29**:36–45.

Liu L, Pan A, Spofford C *et al.* An evolutionary metabolic engineering approach for enhancing lipogenesis in *Yarrowia lipolytica*. *Metab Eng* 2015c;**29**:36–45.

Liu L, Redden H, Alper HS. Frontiers of yeast metabolic engineering: Diversifying beyond ethanol and *Saccharomyces*. *Curr Opin Biotechnol* 2013;**24**:1023–30.

Liu YG, Whittier RF. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 1995;**25**:674–81.

Long TM, Su YK, Headman J *et al.* Cofermentation of glucose, xylose, and cellobiose by the beetle-associated yeast *Spathaspora passalidarum*. *Appl Environ Microbiol* 2012;**78**:5492–500.

Luterbacher JS, Rand JM, Martin Alonso D *et al.* Nonenzymatic Sugar Production from Biomass Using Biomass-Derived γ-Valerolactone. *Science (80-)* 2014:277–80.

Lynd LR. Overview and evaluaton of fuel ethanol from cellulosic biomass: technology, economics, the environment and policy. 1996.

Madden T. The BLAST Sequence Analysis Tool How BLAST Works : The Basics. *October* 2003:1–17.

Markham KA, Palmer CM, Chwatko M *et al.* Rewiring *Yarrowia lipolytica* toward triacetic acid lactone for materials generation . *Proc Natl Acad Sci* 2018;**115**:2096–101.

Markham KA, Vazquez S, Alper HS. High-efficiency transformation of *Yarrowia lipolytica* using electroporation. *FEMS Yeast Res* 2018, DOI: 10.1093/femsyr/foy081.

Marriott PE, Gómez LD, Mcqueen-Mason SJ. Unlocking the potential of lignocellulosic biomass through plant science. *New Phytol* 2016, DOI: 10.1111/nph.13684.

Matsushika A, Inoue H, Kodaki T *et al.* Ethanol production from xylose in engineered

Saccharomyces cerevisiae strains: Current state and perspectives. *Appl Microbiol Biotechnol* 2009;**84**:37–53.

Matthäus F, Ketelhot M, Gatter M *et al.* Production of lycopene in the non-carotenoid-producing yeast *Yarrowia lipolytica*. *Appl Environ Microbiol* 2014, DOI: 10.1128/aem.03167-13.

McMichael P. Agrofuels in the food regime. *J Peasant Stud* 2010, DOI: 10.1080/03066150.2010.512450.

Mcqueen-mason SJ. Tansley review Unlocking the potential of lignocellulosic biomass through plant science. 2015, DOI: 10.1111/nph.13684.

Meacham F, Boffelli D, Dhahbi J *et al.* Identification and correction of systematic error in high-throughput sequence data. *BMC Bioinformatics* 2011;**12**:451.

Milano J, Ong HC, Masjuki HH *et al.* Microalgae biofuels as an alternative to fossil fuel for power generation. *Renew Sustain Energy Rev* 2016;**58**:180–97.

Mol APJ. Environmental authorities and biofuel controversies. *Env Polit* 2010, DOI: 10.1080/09644010903396085.

Mumberg D, Müller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 1995;**156**:119–22.

Naik SN, Goud V V., Rout PK *et al.* Production of first and second generation biofuels: A comprehensive review. *Renew Sustain Energy Rev* 2010;**14**:578–97.

Narisawa T, Fukaura Y, Yazawa K *et al.* Colon cancer prevention with a small amount of dietary perilla oil high in alpha-linolenic acid in an animal model. *Cancer* 1994, DOI: 10.1002/1097-0142(19940415)73:8<2069::AID-CNCR2820730810>3.0.CO;2-1.

Nicaud J-M, Madzak C, van den Broek P *et al.* Protein expression and secretion in the yeast *Yarrowia lipolytica*. *FEMS Yeast Res* 2002;**2**:371–9.

Nijland JG, Shin HY, de Jong RM *et al.* Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2014;**7**:168.

Ogata T, Iwashita Y, Kawada T. Construction of a brewing yeast expressing the

glucoamylase gene STA1 by mating. *J Inst Brew* 2017;**123**:66–9.

Palmer CM, Alper HS. Expanding the Chemical Palette of Industrial Microbes: Metabolic Engineering for Type III PKS-Derived Polyketides. *Biotechnol J* 2019, DOI: 10.1002/biot.201700463.

Pan LX, Yang DF, Shao L *et al.* Isolation of the oleaginous yeasts from the soil and studies of their lipid-producing capacities. *Food Technol Biotechnol* 2009;**47**:215–20.

Parreiras LS, Breuer RJ, Narasimhan RA *et al.* Engineering and two-stage evolution of a lignocellulosic hydrolysate-tolerant *Saccharomyces cerevisiae* strain for anaerobic fermentation of xylose from AFEX pretreated corn stover. *PLoS One* 2014;**9**, DOI: 10.1371/journal.pone.0107499.

Pauly M, Keegstra K. Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J* 2008;**54**:559–68.

Przybysz Buzała K, Kalinowska H, Przybysz P *et al.* Conversion of various types of lignocellulosic biomass to fermentable sugars using kraft pulping and enzymatic hydrolysis. *Wood Sci Technol* 2017, DOI: 10.1007/s00226-017-0916-7.

Qiao K, Imam Abidi SH, Liu H *et al.* Engineering lipid overproduction in the oleaginous yeast *Yarrowia lipolytica*. *Metab Eng* 2015;**29**:56–65.

Qiao K, Wasylenko TM, Zhou K *et al.* Lipid production in *Yarrowia lipolytica* is maximized by engineering cytosolic redox metabolism. *Nat Biotechnol* 2017;**35**:173–7.

Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–2.

Quinn JC, Davis R. The potentials and challenges of algae based biofuels: A review of the techno-economic, life cycle, and resource assessment modeling. *Bioresour Technol* 2015;**184**:444–52.

Reider Apel A, Ouellet M, Szmidt-Middleton H *et al.* Evolved hexose transporter enhances xylose uptake and glucose/xylose co-utilization in *Saccharomyces cerevisiae*. *Sci Rep* 2016;**6**:19512.

Revuelta JL, Ledesma-Amaro R, Lozano-Martinez P *et al.* Bioproduction of riboflavin: a bright yellow history. *J Ind Microbiol Biotechnol* 2017, DOI: 10.1007/s10295-016-1842-7.

Robinson JA. Polyketide synthase complexes: their structure and function in antibiotic biosynthesis. *Philos Trans R Soc Lond B Biol Sci* 1991, DOI: 10.1098/rstb.1991.0038.

Rodriguez-Leyva D, Bassett CMC, McCullough R *et al.* The cardiovascular effects of flaxseed and its omega-3 fatty acid, alpha-linolenic acid. *Can J Cardiol* 2010, DOI: 10.1016/S0828-282X(10)70455-4.

Rodriguez GM, Hussain MS, Gambill L *et al.* Engineering xylose utilization in *Yarrowia lipolytica* by understanding its cryptic xylose pathway. *Biotechnol Biofuels* 2016;**9**:1–15.

Roostita R, Fleet GH. The occurrence and growth of yeasts in Camembert and Blue-veined cheeses. *Int J Food Microbiol* 1996, DOI: 10.1016/0168-1605(95)00018-6.

Ruiz-Herrera J, Sentandreu R. Different effectors of dimorphism in *Yarrowia lipolytica*. *Arch Microbiol* 2002;**178**:477–83.

Ryu S, Hipp J, Trinh CT. Activating and Elucidating Complex Sugar Metabolism in *Yarrowia lipolytica*. *Appl Env Microbiol* 2015:1–48.

Salusjärvi L, Kankainen M, Soliymani R *et al.* Regulation of xylose metabolism in recombinant *Saccharomyces cerevisiae*. *Microb Cell Fact* 2008;**7**:18.

Salusjärvi L, Pitkänen J-P, Aristidou A *et al.* Transcription analysis of recombinant *Saccharomyces cerevisiae* reveals novel responses to xylose. *Appl Biochem Biotechnol* 2006;**128**:237–61.

Sànchez Nogué V, Karhumaa K. Xylose fermentation as a challenge for commercialization of lignocellulosic fuels and chemicals. *Biotechnol Lett* 2015;**37**:761–72.

Sanchez RG, Solodovnikova N, Wendland J. Breeding of lager yeast with *Saccharomyces cerevisiae* improves stress resistance and fermentation performance. *Yeast* 2012;**29**:343–55.

Scalcinati G, Otero JM, Van Vleet JRH *et al.* Evolutionary engineering of *Saccharomyces cerevisiae* for efficient aerobic xylose consumption. *FEMS Yeast Res* 2012;**12**:582–97.

Schwechheimer SK, Park EY, Revuelta JL *et al.* Biotechnology of riboflavin. *Appl Microbiol Biotechnol* 2016, DOI: 10.1007/s00253-015-7256-z.

Sedlak M, Ho NWY. Characterization of the effectiveness of hexose transporters for transporting xylose during glucose and xylose co-fermentation by a recombinant Saccharomyces yeast. *Yeast* 2004;**21**:671–84.

Shen Y, Chen X, Peng B *et al.* An efficient xylose-fermenting recombinant *Saccharomyces cerevisiae* strain obtained through adaptive evolution and its global transcription profile. *Appl Microbiol Biotechnol* 2012;**96**:1079–91.

Shin HY, Nijland JG, de Waal PP *et al.* An engineered cryptic Hxt11 sugar transporter facilitates glucose–xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2015;**8**:176.

Da Silva ASA, Teixeira RSS, Endo T *et al.* Continuous pretreatment of sugarcane bagasse at high loading in an ionic liquid using a twin-screw extruder. *Green Chem* 2013, DOI: 10.1039/c3gc40352a.

Snoek T, Picca Nicolino M, Van Den Bremt S *et al.* Large-scale robot-assisted genome shuffling yields industrial *Saccharomyces cerevisiae* yeasts with increased ethanol tolerance. *Biotechnol Biofuels* 2015;**8**:1–19.

Solis-Escalante D, van den Broek M, Kuijpers NGA *et al.* The genome sequence of the popular hexose-transport-deficient *Saccharomyces cerevisiae* strain EBY.VW4000 reveals LoxP/Cre-induced translocations and gene loss. *FEMS Yeast Res* 2015;**15**:1–12.

Spagnuolo M, Hussain MS, Gambill L *et al.* Alternative substrate metabolism in *Yarrowia lipolytica*. *Front Microbiol* 2018, DOI: 10.3389/fmicb.2018.01077.

Stark AH, Crawford MA, Reifen R. Update on alpha-linolenic acid. *Nutr Rev* 2008, DOI: 10.1111/j.1753-4887.2008.00040.x.

Steensels J, Meersman E, Snoek T *et al.* Large-scale selection and breeding to generate

industrial yeasts with superior aroma production. *Appl Environ Microbiol* 2014a;**80**:6965–75.

Steensels J, Snoek T, Meersman E *et al.* Improving industrial yeast strains: Exploiting natural and artificial diversity. *FEMS Microbiol Rev* 2014b;**38**:947–95.

Subtil T, Boles E. Competition between pentoses and glucose during uptake and catabolism in recombinant *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2012;**5**:14.

Taheripour F, Hertel TW, Tyner WE *et al.* Biofuels and their by-products: Global economic and environmental implications. *Biomass and Bioenergy* 2010, DOI: 10.1016/j.biombioe.2009.10.017.

Tai M. Metabolic Engineering of oleaginous yeast for the production of biofuels. *Diss - MIT* 2012.

Takács E, Wojnárovits L, Földváry C *et al.* Effect of combined gamma-irradiation and alkali treatment on cotton-cellulose. *Radiation Physics and Chemistry*. 2000.

Van Vleet JH, Jeffries TW. Yeast metabolic engineering for hemicellulosic ethanol production. *Curr Opin Biotechnol* 2009;**20**:300–6.

Wagner JM, Alper HS. Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances. *Fungal Genet Biol* 2015:1–11.

Wagner JM, Liu L, Yuan SF *et al.* A comparative analysis of single cell and droplet-based FACS for improving production phenotypes: Riboflavin overproduction in *Yarrowia lipolytica*. *Metab Eng* 2018;**47**:346–56.

Wagner JM, Williams E V., Alper HS. Developing a piggyBac transposon system and compatible selection markers for insertional mutagenesis and genome engineering in *Yarrowia lipolytica*. *Biotechnol J* 2018, DOI: 10.1002/biot.201800022.

Walfridsson M, Hallborn J, Penttilä M *et al.* Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. Saccharomyces. *Appl Environ Microbiol* 1995;**61**:4184–90.

Wang M, Yu C, Zhao H. Directed evolution of xylose specific transporters to facilitate

glucose-xylose co-utilization. *Biotechnol Bioeng* 2015;**9999**:1–8.

Wang PY, Shopsis C, Schneider H. Fermentation of a pentose by yeasts. *Biochem Biophys Res Commun* 1980;**94**:248–54.

Wasylenko TM, Ahn WS, Stephanopoulos G. The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from glucose in *Yarrowia lipolytica*. *Metab Eng* 2015a, DOI: 10.1016/j.ymben.2015.02.007.

Wasylenko TM, Ahn WS, Stephanopoulos G. The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from glucose in *Yarrowia lipolytica*. *Metab Eng* 2015b;**30**:27–39.

Weber H, Kurischko C, Barth G. Mating in the alkane-utilizing yeast *Yarrowia lipolytica*. 1988;**4**:229–40.

Wieczorke R, Krampe S, Weierstall T *et al.* Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in *Saccharomyces cerevisiae*. *FEBS Lett* 1999;**464**:123–8.

Xue S, Jones AD, Sousa L *et al.* Water-soluble phenolic compounds produced from extractive ammonia pretreatment exerted binary inhibitory effects on yeast fermentation using synthetic hydrolysate. *PLoS One* 2018, DOI: 10.1371/journal.pone.0194012.

Young E, Lee S-M, Alper H. Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnol Biofuels* 2010;**3**:24.

Young E, Poucher A, Comer A *et al.* Functional survey for heterologous sugar transport proteins, using *Saccharomyces cerevisiae* as a host. *Appl Environ Microbiol* 2011;**77**:3311–9.

Young EM, Comer AD, Huang H *et al.* A molecular transporter engineering approach to improving xylose catabolism in *Saccharomyces cerevisiae*. *Metab Eng* 2012;**14**:401–11.

Young EM, Tong A, Bui H *et al.* Rewiring yeast sugar transporter preference through modifying a conserved protein motif. *Proc Natl Acad Sci U S A* 2014;**111**:131–6.

Zhang YHP, Lynd LR. A functionally based model for hydrolysis of cellulose by fungal

cellulase. *Biotechnol Bioeng* 2006, DOI: 10.1002/bit.20906.

Zheng J, Rehmann L. Extrusion pretreatment of lignocellulosic biomass: A Review. *Int J Mol Sci* 2014;**15**:18967–84.